

Adaptive Algorithms in Finite Form

Ivan Tyukin^{*}Danil Prokhorov[†]and Cees van Leeuwen^{*}

Abstract

We propose a new method for the design of adaptation algorithms that guarantees a certain prescribed level of performance and applicable to systems with nonconvex parameterization. The main idea behind the method is two-fold. First, we augment the tuning error function and design the adaptation scheme in the form of ordinary differential equations. The resulting augmentation is allowed to depend on state derivatives. Second, we find a suitable realization of the designed adaptation scheme in an algebraic-integral form. Due to their explicit dependence on the state of the original system, such adaptation schemes are referred to as adaptive algorithms in *finite form*, in contrast to (conventional) algorithms in differential form. Sufficient conditions for the existence of finite form realizations are proposed. It is shown that our method to design algorithms in finite form is applicable to a broad class of nonlinear systems including systems with nonconvex parameterization and low-triangular systems.

Keywords: adaptive systems, performance, nonlinear parameterization, finite form algorithms

1 Introduction

Significant progress in adaptive control theory has been made in the areas of linear and nonlinear systems [3, 13, 20, 26], plants with relative degree greater than one [10, 19, 18], and systems with nonconvex parameterization [17, 15]. However, there is still room for further developments, as there are important unresolved problems regarding the issue of performance, especially in the presence of nonconvex parameterization.

As expressed by asymptotic stability of adaptive systems [26], robustness, and good transient behavior [14], suitable performance can be proven under the requirement of persistent excitation. As it is generally observed in practice, insufficient excitation results in absence of asymptotic stability and, as its consequence, in poor parameter convergence, sensitivity to small disturbances and poor transient performance. A suitable performance criterion is needed to assure the efficiency and quality of the system. As substitutes for performance criteria, most of the available results in direct adaptive control without restrictive persistent excitation requirements limit themselves to L_2 and L_∞ ¹ norm bounds of the tracking errors. For more sophisticated performance measures like the LQ criterion, some results are available [7, 6]. These results, however, deal either with too narrow a class of uncertain systems [7] or present only a comparison between adaptive and robust backstepping [6] without suggesting new adaptation schemes. On the other hand, when improvement *heuristics* are suggested like in [21], no exact performance criterion is provided that can explicitly be computed *a-priori*, except probably the bounds on L_2 and L_∞ norms for the tracking errors.

^{*}Laboratory for Perceptual Dynamics, RIKEN (Institute for Physical and Chemical Research) Brain Science Institute, 2-1, Hirosawa, Wako-shi, Saitama, 351-0198, Japan, e-mail: {tyukinivan,ceesvl}@brain.riken.jp

[†]Ford Research Laboratory, Dearborn, MI, 48121, USA, e-mail: dprokhor@ford.com

¹Function $\nu : R_+ \rightarrow R$ is said to belong to L_2 iff $L_2(\nu) = \int_0^\infty \nu^2(\tau) d\tau < \infty$. The value $\sqrt{L_2(\nu)}$ stands for the L_2 norm of $\nu(t)$. Function $\nu : R_+ \rightarrow R$ belongs to and L_∞ iff $L_\infty(\nu) = \sup_{t \geq 0} \|\nu(t)\| < \infty$, where $\|\cdot\|$ is the Euclidean norm. The value of $L_\infty(\nu)$ stands for the L_∞ norm of $\nu(t)$.

Another unresolved issue in conventional adaptive control theory is nonconvex parameterization of the plant model. The available approaches encourage to compensate for the nonlinearity by using an additional damping term [17, 15, 22]. These techniques guarantee existence of a solution to the control problem, but are limited in terms of their practical value because of their high-gain nature. Moreover, none of these approaches provides a performance measure beyond L_∞ or L_2 norm bounds with respect to the state or tracking error.

An important impediment to solving these two major problems at once, we believe, is the lack of sufficient information in the conventional adaptive schemes. One way to provide the algorithms with extra information is to augment the tuning errors. This idea is inherent to both Morse's adaptive controllers [19] and those based on Kreisselmeier's observers [10] when dealing with plants with relative degree greater than one. These augmented errors then are used in conventional gradient schemes

$$\dot{\hat{\theta}} = -\Gamma\psi(\mathbf{x}, t)\mathcal{A}(\mathbf{x}, \hat{\theta}, t), \quad (1)$$

where $\mathbf{x} \in R^n$ is a state (or output) vector, $\hat{\theta} \in R^d$ is a vector of controller parameters, $\mathcal{A}(\mathbf{x}, \hat{\theta})$ is an operator that depends on the particular problem at hand, $\psi(\mathbf{x}, t)$ is the error function and gain $\Gamma > 0$. Improved performance of these augmented controllers for Morse's high-order tuners is reported in [2, 23]. In these papers the plants are assumed to be linear, but no additional criteria are provided except for L_2 and L_∞ norm bounds with respect to the tracking errors.

The above-mentioned limitations of performance and applicability motivate us to search for a new augmentation that uses additional information about the system dynamics, such as, e.g., state derivatives. As a result, new properties in the system can be created. On the other hand, we wish to find physically realizable algorithms that do not require measurements of any unknown signals, derivatives, or parameters. In order to meet these seemingly contradictory requirements we propose to extend conventional algorithms (1) as follows:

$$\hat{\theta}(\mathbf{x}(t), t) = \hat{\theta}_P(\mathbf{x}, t) + \hat{\theta}_I(t); \quad \dot{\hat{\theta}}_I = \mathcal{A}_2(\mathbf{x}, \hat{\theta}, t), \quad \hat{\theta}_P(\mathbf{x}, t) = \mathcal{A}_1(\mathbf{x}, t). \quad (2)$$

Notice that if functions $\hat{\theta}(\mathbf{x}, t)$ are written in differential form (1), they may depend on unknown parameters and unmeasured signals, e.g., state derivatives. Thus, the equivalent description of adaptive algorithm (2) in differential form may produce an augmentation that is in fact derivative-dependent, thereby providing the algorithm with more information about the plant uncertainties. These observations lead to quite unexpected consequences. Instead of restricting the design procedure to the algorithms given in differential form (1), it becomes possible to design adaptation algorithms in two steps. First, search for the desired augmentation to obtain the required adaptive control properties. At this stage it does not matter whether the augmentation is uncertainty-dependent or not. Once a suitable tuning error is chosen, second: find a realization of the algorithm in the form of integral-algebraic equation (2). Such a realization will be termed *algorithm in finite form*.

According to our knowledge, algorithms (2) have been introduced for direct adaptive control of nonlinear systems in 1986 in [4] and then were reintroduced later in [24, 1]. Their distinctive performance properties and extended applicability, however, were not appreciated at the time. The efficiency of the proposed algorithms in

[4, 5] was limited by restrictive pseudo-gradient assumptions on $\hat{\theta}_P(\mathbf{x}, t)$ (for the details see [4, 5]). Nevertheless, it has been reported recently that algorithms (2) may be able to deal with nonconvex parameterization (see for example [30], Lemma 1, p. 558; [24, 28]) and guarantee improved transient performance [27, 28]. Some preliminary analysis of the distinctive properties of algorithms (2) is available, for example, in [28, 29]. Additional support and motivation for algorithms (2) can be found in [1], where the authors introduced their adaptation schemes from immersion and invariance principles. However, the main problem, with the current study of algorithms (2) is that there is no systematic method that allows us to design these algorithms with guaranteed improvements in performance and, at the same time, achieve applicability to systems with nonlinear parameterization for a sufficiently broad class of nonlinear dynamical systems.

In our present work we suggest a new method to design adaptive algorithms in finite form (2) that guarantee improved performance and in addition are applicable to a class of nonlinearly parameterized plants. The method is systematic and is based on two fundamental ideas in adaptive control theory: *augmentation* of the error and *embedding* the original system dynamics into one of a higher order. These ideas are embodied in two independent stages of the design. The first stage is *augmentation* of the tuning error for algorithms in the conventional differential form (1) in order to ensure improved performance and extended applicability of these algorithms. The resulting augmentation may not necessarily be independent on the uncertainties or time-derivatives of the state vector. The resulting augmentation, however, should guarantee certain desired properties of the system.

Based on the augmentation (possibly, derivative-dependent) obtained in the first stage, the second stage of the design method should be to find functions $\mathcal{A}_1(\mathbf{x}, t)$, $\mathcal{A}_2(\mathbf{x}, \hat{\theta}, t)$ which guarantee that algorithms (2) realize the desired adaptation scheme. We show that this problem may require finding a solution of a system of partial differential equations. It is well-known that such a solution may not exist in general. To avoid this problem, we consider several special cases of plant models, with their structures satisfying sufficient conditions for the existence of a solution. As soon as these basic structures are found, we *embed* the original system into a system of higher order for which the solution is known to exist. The embedding is to be made in such a way that the extended system belongs to one of already established basic classes that guarantee existence of the solution to the realization problem. With embedding we shall be able to obtain adaptation schemes that guarantee not only square integrability of the error but also integrability of its first derivatives as well as square integrability of control efforts injected into the system due to the parametric uncertainties. In addition, we provide the conditions for which the decrease of the parametric uncertainties and exponential convergence into a neighborhood of the target manifold are guaranteed without the restrictive assumption of persistent excitation. Last but not the least, our new adaptive schemes can also be applied to systems with nonlinear parameterization.

The layout of the paper is as follows. In Section 2 we specify the class of nonlinear dynamical systems under consideration and select the desired augmentation. Section 3 contains the main results of the paper. In Section 4 we present an example of the design and results of computer simulations. Section 5 concludes the paper.

2 Problem Formulation and Preliminary Results

Let the plant mathematical model be given as follows:

$$\begin{aligned}\dot{x}_i &= f_i(\mathbf{x}) + g_i(\mathbf{x})u, \quad i = 1, \dots, m \\ \dot{x}_j &= f_j(\mathbf{x}) + \nu_{j-m}(\mathbf{x}, \boldsymbol{\theta}) + g_j(\mathbf{x})u, \quad j = m+1, \dots, n,\end{aligned}\tag{3}$$

where $\mathbf{x} \in R^n$ is a state vector, $f_i, g_i : R^n \rightarrow R$, $f_i, g_i \in C^1$, $\boldsymbol{\theta} \in \Omega_\theta \subset R^d$ is a vector of unknown parameters, $\nu_i : R^n \times R^d \rightarrow R$, $\nu_i \in C^1$, u is a control input. Let us define functions $\mathbf{f}(\cdot)$, $\mathbf{g}(\cdot)$, $\boldsymbol{\nu}(\cdot, \cdot)$:

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_n(\mathbf{x}))^T, \quad \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_n(\mathbf{x}))^T, \quad \boldsymbol{\nu}(\mathbf{x}, \boldsymbol{\theta}) = (\nu_1(\mathbf{x}, \boldsymbol{\theta}), \dots, \nu_{n-m}(\mathbf{x}, \boldsymbol{\theta}))^T$$

It will be useful sometimes to think of state vector $\mathbf{x} \in \mathcal{L} \subseteq R^n$ as $\mathbf{x} = \mathbf{x}_1 \oplus \mathbf{x}_2$, $\mathbf{x}_1 = (x_1, x_2, \dots, x_m)^T$, $\mathbf{x}_2 = (x_{m+1}, x_{m+2}, \dots, x_n)^T$, where symbol \oplus denotes concatenation of two vectors $\mathbf{x}_1 \in \mathcal{L}_1 \subseteq R^m$, $\mathbf{x}_2 \in \mathcal{L}_2 \subseteq R^{n-m}$ and \mathcal{L} , \mathcal{L}_1 , \mathcal{L}_2 are linear spaces. The time-derivative of \mathbf{x}_1 is independent on $\boldsymbol{\theta}$, whereas the time-derivative of vector \mathbf{x}_2 depends on unknown parameters $\boldsymbol{\theta}$ explicitly. Therefore we refer to the spaces \mathcal{L}_1 and \mathcal{L}_2 as *uncertainty-independent* and *uncertainty-dependent partitions* of system (3), respectively. To denote the right-hand sides of the partitioned system, we use the following notations: $\mathbf{f}_1 = (f_1, \dots, f_m)^T$, $\mathbf{f}_2 = (f_{m+1}, \dots, f_n)^T$, $\mathbf{g}_1 = (g_1, \dots, g_m)^T$, $\mathbf{g}_2 = (g_{m+1}, \dots, g_n)^T$. Hence, the partitioned system can be written as follows:

$$\dot{\mathbf{x}}_1 = \mathbf{f}_1(\mathbf{x}) + \mathbf{g}_1(\mathbf{x})u, \quad \dot{\mathbf{x}}_2 = \mathbf{f}_2(\mathbf{x}) + \boldsymbol{\nu}(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{g}_2(\mathbf{x})u.\tag{4}$$

In analogy with the definition of independence of a function with respect to the components x_i of its argument \mathbf{x} , we would like to define a notion of independence of the function with respect to the partition. Let $\mathbf{x} \in \mathcal{L} = \mathcal{L}_1 \oplus \mathcal{L}_2$, and let function $\omega(\mathbf{x}) : R^n \rightarrow R^n$ be differentiable for any $\mathbf{x} \in R^n$. Function $\omega(\mathbf{x})$ is said to be *independent on partition \mathcal{L}_2* iff $\partial\omega(\mathbf{x}_1 \oplus \mathbf{x}_2)/\partial\mathbf{x}_2 = 0$. We would also like to extend the standard definition of the Lie derivatives to the partitioned system. Given the following partition $\mathbf{x} = \mathbf{x}_1 \oplus \mathbf{x}_2 \oplus \dots \oplus \mathbf{x}_r$, we denote by symbol $L_{\mathbf{f}_i}\psi(\mathbf{x})$ the following derivatives $(\partial\psi(\mathbf{x})/\partial\mathbf{x}_i) \mathbf{f}_i(\mathbf{x}) = L_{\mathbf{f}_i}\psi(\mathbf{x})$, where $\mathbf{f}_i(\mathbf{x})$ stands for the corresponding vector-function in $\mathbf{f}(\mathbf{x}) = \mathbf{f}_1(\mathbf{x}) \oplus \mathbf{f}_2(\mathbf{x}) \oplus \dots \oplus \mathbf{f}_r(\mathbf{x})$. If function $\boldsymbol{\psi}(\mathbf{x}) = (\psi_1(\mathbf{x}), \dots, \psi_n(\mathbf{x}))^T$ is a vector-function then symbol $L_{\mathbf{f}_i}\boldsymbol{\psi}(\mathbf{x})$ denotes the following vector: $L_{\mathbf{f}_i}\boldsymbol{\psi}(\mathbf{x}) = (L_{\mathbf{f}_i}\psi_1(\mathbf{x}), \dots, L_{\mathbf{f}_i}\psi_n(\mathbf{x}))^T$.

As in [30, 25], we define the control goal as reaching asymptotically a target manifold. We assume that the target manifold can be given by the following equality $\psi(\mathbf{x}, t) = 0$, where $\psi : R^n \times R \rightarrow R$, $\psi(\mathbf{x}, t) \in C^1$. Additional restrictions on the function $\psi(\mathbf{x}, t)$ are formulated in Assumptions 1, 2.

Assumption 1 (Boundedness of the Solutions) *Function $\psi(\mathbf{x}, t)$ is such that for any $\delta > 0$ there exists a function $\varepsilon : R_+ \rightarrow R_+$ such that $|\psi(\mathbf{x}, t)| \leq \delta \Rightarrow \|\mathbf{x}\| \leq \varepsilon(\delta)$ along system (3) solutions.*

Assumption 1 simply states that any trajectory of system (3) belonging to a neighborhood of the target manifold $\psi(\mathbf{x}, t)$ is bounded. Clearly, most of the common goal criteria used in adaptive control satisfy this property, for example, positive-definite functions $\psi(\mathbf{x})$ for nonlinear systems and quadratic forms for linear ones. In general,

however, in order to show the boundedness of \mathbf{x} it is not necessary for the function $\psi(\mathbf{x})$ to be positive definite.

As an illustration, consider the following system:

$$\begin{aligned}\dot{x}_i &= f_i(x_1, \dots, x_i) + x_{i+1}, \quad i = \{1, \dots, n-1\} \\ \dot{x}_n &= f_n(x_1, \dots, x_n) + \nu(\mathbf{x}, \boldsymbol{\theta}) + u.\end{aligned}\tag{5}$$

Let $\psi(\mathbf{x}) = x_n - p(x_1, \dots, x_{n-1}) + f_{n-1}(x_1, \dots, x_{n-1})$, $f_i(\cdot), p(\cdot) \in C^0$ and furthermore, let the system

$$\begin{aligned}\dot{x}_i &= f_i(x_1, \dots, x_i) + x_{i+1}, \quad i = \{1, \dots, n-2\} \\ \dot{x}_{n-1} &= p(x_1, \dots, x_{n-1}) + v\end{aligned}\tag{6}$$

state be bounded for any $v \in L_\infty$ (i.e. system (6) has the *bounded input - bounded state* property). Then for the system of equations (5), it is sufficient that system (6) is input-to-state stable with respect to input v to satisfy Assumption 1² with $\psi(\mathbf{x}) = x_n - p(x_1, \dots, x_{n-1}) + f_{n-1}(x_1, \dots, x_{n-1})$.

Assumption 2 (Regularity) *For any $\mathbf{x} \in R^n$ and $t > 0$ functions $\psi(\mathbf{x}, t)$ and $\mathbf{g}(\mathbf{x})$ satisfy the following inequality $|L_{\mathbf{g}}\psi(\mathbf{x}, t)| > \delta_1 > 0$.*

Assumption 2 ensures the existence of feedback that transforms the original system into that of the error model with respect to the variable $\psi(\mathbf{x}, t)$. Let Assumption 2 hold; consider

$$\dot{\psi} = L_{\mathbf{f}}\psi(\mathbf{x}, t) + L_{\nu(\mathbf{x}, \boldsymbol{\theta})}\psi(\mathbf{x}, t) + (L_{\mathbf{g}}\psi(\mathbf{x}, t))u + \partial\psi(\mathbf{x}, t)/\partial t.\tag{7}$$

Because of Assumption 2 there exists the control input

$$u(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) = (L_{\mathbf{g}(\mathbf{x})}\psi(\mathbf{x}, t))^{-1} \left(-\varphi(\psi) - L_{\mathbf{f}}\psi(\mathbf{x}, t) - L_{\nu(\mathbf{x}, \hat{\boldsymbol{\theta}})}\psi(\mathbf{x}, t) - \partial\psi(\mathbf{x}, t)/\partial t \right),\tag{8}$$

where $\hat{\boldsymbol{\theta}} \in \Omega_{\hat{\boldsymbol{\theta}}} \subset R^d$ – is a vector of controller parameters that transforms (7) into

$$\dot{\psi} = -\varphi(\psi) + z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)\tag{9}$$

where $z(\mathbf{x}, \boldsymbol{\theta}, t) = L_{\nu(\mathbf{x}, \boldsymbol{\theta})}\psi(\mathbf{x}, t)$. Let the closed loop system satisfy some additional requirements:

Assumption 3 (Certainty Equivalence) *For any $\boldsymbol{\theta} \in \Omega_{\boldsymbol{\theta}}$ there exists $\hat{\boldsymbol{\theta}}^* \in \Omega_{\hat{\boldsymbol{\theta}}} \subset R^d$, such that for all $\mathbf{x} \in R^n$, $t \in R_+$ the following equivalence holds*

$$\frac{\partial\psi(\mathbf{x}, t)}{\partial \mathbf{x}}[\mathbf{f}(\mathbf{x}) + \nu(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{g}(\mathbf{x})u(\mathbf{x}, \hat{\boldsymbol{\theta}}^*, t)] + \frac{\partial\psi(\mathbf{x}, t)}{\partial t} + \varphi(\psi) = \dot{\psi}(\mathbf{x}, \boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*, t) + \varphi(\psi) = 0.\tag{10}$$

It is clear that if Assumption 2 holds then Assumption 3 is automatically satisfied. According to Assumptions 2 and 3, it follows that $z(\mathbf{x}, \hat{\boldsymbol{\theta}}^*, t) = z(\mathbf{x}, \boldsymbol{\theta}, t)$ for any $\mathbf{x} \in R^n$ and time $t > 0$.

Assumption 4 (Stability of the Target Dynamics) *Function $\varphi(\psi)$ in (10) satisfies*

$$\varphi(\psi) \in C^0, \quad \varphi(\psi)\psi > 0 \quad \forall \psi \neq 0, \quad \lim_{\psi \rightarrow \infty} \int_0^\psi \varphi(\varsigma) d\varsigma = \infty.\tag{11}$$

²For plants with uncertainties in functions $f_i(\cdot)$, the time-dependent target manifold $\psi(\mathbf{x}, t) = 0$ may be required for control design. This will be clarified later in the proof of Theorem 4 and in the examples section.

Assumption 5 (Monotonicity and Linear Growth Rate in Parameters) *There exists function $\alpha(\mathbf{x}, t) : R^n \times R \rightarrow R^d$ such that $(z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}^*, t))(\alpha(\mathbf{x}, t)^T(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)) > 0 \forall z(\mathbf{x}, \hat{\boldsymbol{\theta}}^*, t) \neq z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)$. Furthermore, the following holds: $|z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}^*, t)| \leq D|\alpha(\mathbf{x}, t)^T(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)|$, $D > 0$.*

Assumption 3 (certainty equivalence or matching condition) simply states that for every unknown $\boldsymbol{\theta}^* \in \Omega_\theta$ there exists a vector of controller parameters $\hat{\boldsymbol{\theta}}^*(\boldsymbol{\theta}^*) \in \Omega_{\hat{\boldsymbol{\theta}}}$ such that the system dynamics with this control function satisfies the following equation $\dot{\psi} = -\varphi(\psi)$. Assumption 4 specifies the properties of function $\varphi(\psi)$, thus stipulating asymptotic stability of manifold $\psi(\mathbf{x}, t) = 0$ for $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}^*$ and ensuring unbounded growth of integral $\int_0^\psi \varphi(\varsigma) d\varsigma$ as $\psi \rightarrow \infty$. Assumption 5 is given to specify the admissible nonlinear parameterization of the controller. For linearly parameterized plants this assumption is automatically satisfied. Sometimes we will further restrict the nonlinear in parameter functions by:

Assumption 6 *There exists a positive constant $D_1 > 0$ such that for any $\mathbf{x}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}^*, t > 0$ the following inequality holds: $|z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}^*, t)| \geq D_1|\alpha(\mathbf{x}, t)^T(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)|$.*

Throughout the paper we will assume that functions $\alpha(\mathbf{x}, t)$ and $u(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)$ are both bounded in t . For the sake of convenience and if not stated otherwise we will also assume that functions $\alpha(\mathbf{x}, t), \psi(\mathbf{x}, t)$ are differentiable as many times as necessary if differentiation is required to design the algorithm. In addition, we will use the term "smooth functions" to denote those functions that belong to C^∞ . Though it is not necessary at all for us to require existence of infinitely many derivatives of the functions that we refer to as smooth in the paper, this notational agreement will free the presentation of numerous insignificant details in the formulations. Furthermore, along with already defined L_2 and L_∞ norms we will use the following notation: $\|\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}(t)\|_{\Gamma^{-1}}^2 = (\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}(t))^T \Gamma^{-1} (\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}(t))$.

As mentioned in the introduction, we propose to design the adaptive algorithms in two steps: 1) search for the suitable augmentation ensuring the desired properties of the control system, and 2) find the appropriate realization of this algorithm in finite form. Therefore we start with the choice of tuning errors $\tilde{\psi}(\mathbf{x}, t)$ and operators $\mathcal{A}(\mathbf{x}, t)$ for the class of algorithms given by formula (1): $\dot{\hat{\boldsymbol{\theta}}} = \Gamma \tilde{\psi}(\mathbf{x}, t) \mathcal{A}(\mathbf{x}, \hat{\boldsymbol{\theta}})$. As a candidate for the augmented error $\tilde{\psi}(\mathbf{x}, t)$ we select the following $\tilde{\psi}(\mathbf{x}, t) = \dot{\psi} + \psi(\mathbf{x}, t)$. It has been shown in [25, 30] that the algorithm

$$\dot{\hat{\boldsymbol{\theta}}} = \Gamma(\varphi(\psi) + \dot{\psi})\alpha(\mathbf{x}, t), \quad \Gamma > 0 \quad (12)$$

with control (8) guarantee that $\psi(\mathbf{x}(t), t) \rightarrow 0$ as $t \rightarrow \infty$ for the closed loop system

$$\begin{aligned} \dot{\mathbf{x}}_1 &= \mathbf{f}_1(\mathbf{x}) + \mathbf{g}_1(\mathbf{x})u, & \dot{\mathbf{x}}_2 &= \mathbf{f}_2(\mathbf{x}) + \boldsymbol{\nu}(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{g}_2(\mathbf{x})u \\ \dot{\hat{\boldsymbol{\theta}}} &= \Gamma(\varphi(\psi) + \dot{\psi})\alpha(\mathbf{x}, t), & \Gamma &> 0 \end{aligned} \quad (13)$$

under Assumptions 1, 3 – 5. In addition, it is possible to show that system (13) has better performance than that of the known schemes. This follows from the next theorem (see also Proposition 1 below):

Theorem 1 *Let system (13) be given and Assumptions 2–5 hold. Then for system (13) the following hold:*

P1) $\varphi(\psi(t)) \in L_2$, $\dot{\psi}(t) \in L_2$;

P2) $\|\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}(t)\|_{\Gamma^{-1}}^2$ is non-increasing;

P3) $z((\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}(t), t)) \in L_2$.

Furthermore,

$$\begin{aligned} \|\varphi(\psi)\|_2^2 &\leq 2Q(\psi) + \|\hat{\boldsymbol{\theta}}(0) - \hat{\boldsymbol{\theta}}^*\|_{(2D\Gamma)^{-1}}^2, \quad \|\dot{\psi}\|_2^2 \leq 2Q(\psi) + \|\hat{\boldsymbol{\theta}}(0) - \hat{\boldsymbol{\theta}}^*\|_{(2D\Gamma)^{-1}}^2 \\ \|\psi\|_\infty &\leq \Lambda \left(Q(\psi) + \|\hat{\boldsymbol{\theta}}(0) - \hat{\boldsymbol{\theta}}^*\|_{(4D\Gamma)^{-1}}^2 \right), \end{aligned} \quad (14)$$

where $Q(\psi) = \int_0^{\psi(\mathbf{x}(0), 0)} \varphi(\varsigma) d\varsigma$ and $\Lambda(d) = \max_{|\psi|} \{|\psi| \mid \int_0^{|\psi|} \varphi(\varsigma) d\varsigma = d\}$.

If Assumption 1 is satisfied and function $z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)$ is locally bounded with respect to \mathbf{x} , $\hat{\boldsymbol{\theta}}$ and uniformly bounded with respect to t , then

P4) trajectories of the system are bounded and $\psi(\mathbf{x}(t)) \rightarrow 0$ as $t \rightarrow \infty$;

If in addition functions $\varphi, z(\mathbf{x}, \boldsymbol{\theta}, t) \in C^1$; derivative $\partial z(\mathbf{x}, \boldsymbol{\theta}, t)/\partial t$ is uniformly bounded in t ; function $\boldsymbol{\alpha}(\mathbf{x}, t)$ is locally bounded with respect to \mathbf{x} and uniformly bounded with respect to t , then

P5) $\dot{\psi} \rightarrow 0$ as $t \rightarrow \infty$; $z((\mathbf{x}, \boldsymbol{\theta}^*, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}(t), t)) \rightarrow 0$ as $t \rightarrow \infty$.

The formal proof of the theorem is given in Appendix 2³.

Notice that the function $\varphi(\psi)$ is nonlinear and that its shape influences the L_2 and L_∞ norm bounds for ψ and $\dot{\psi}$. Because of this, according to (14) it is possible to improve the performance of the system with respect to L_2 and L_∞ bounds by varying the function $\varphi(\psi)$. The bounds obtained for the L_∞ norms may be improved further for the case when the function $\varphi(\psi)$ is linear in ψ . This is not too severe a restriction as the choice of function $\varphi(\psi)$ is always up to the designer. Performance characteristics of the system for this case are formulated in Proposition 1.

Proposition 1 (Exponential Convergence) *Let Assumptions 2-5 hold and $\varphi(\psi) = K\psi$, $K > 0$. Then*

P6) *function $\psi(\mathbf{x}(t), t)$ converges exponentially fast into the domain $|\psi(\mathbf{x}(t), t)| \leq 0.5\sqrt{\|\hat{\boldsymbol{\theta}}(0) - \hat{\boldsymbol{\theta}}^*\|_{(K D\Gamma)^{-1}}^2}$. Specifically, the following holds: $|\psi(\mathbf{x}(t), t)| \leq |\psi(\mathbf{x}(0), 0)|e^{-Kt} + 0.5\sqrt{\|\hat{\boldsymbol{\theta}}(0) - \hat{\boldsymbol{\theta}}^*\|_{(K D\Gamma)^{-1}}^2}$*

Furthermore, let Assumption 1 hold, function $\boldsymbol{\alpha}(\mathbf{x}, t)$ be locally bounded with respect to \mathbf{x} and uniformly bounded in t ; for any bounded \mathbf{x} there exist $D_1 > 0$ such that $|z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}^*, t)| \geq D_1|\boldsymbol{\alpha}(\mathbf{x}, t)^T(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)|$, function $\boldsymbol{\alpha}(\mathbf{x}, t)$ is persistently exciting:

$$\exists L > 0, \delta > 0 : \int_t^{t+L} \boldsymbol{\alpha}(\mathbf{x}(\tau), \tau) \boldsymbol{\alpha}(\mathbf{x}(\tau), \tau)^T d\tau \geq \delta I \quad \forall t > 0, \quad (15)$$

where $I \in \mathbb{R}^{d \times d}$ – identity matrix. Then

P7) both $\psi(\mathbf{x}(t), t)$ and $\|\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*\|$ converge exponentially fast to the origin.

It follows from Proposition 1 that if $\varphi(\psi) = K\psi$ then the estimate of the upper bound $\sup_{t \geq t'} |\psi(\mathbf{x}(t), t)|$ as a function of t' for $\psi(\mathbf{x}(t), t)$ in system (13) exponentially converges into the domain determined by the parametric

³Robustness and some other properties of algorithms (12) are discussed in Appendix 1.

uncertainty, the values of controller parameters K , and adaptation gain Γ . Notice that this domain can be made arbitrary small, subject to the choice of the values of K and Γ . The rates of convergence are given by P6). In the case of persistent excitation an even stronger property is established. The system is shown to be *exponentially stable* with respect to the target manifold $\psi(\mathbf{x}, t) = 0$ and point $\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}^*$.

Despite properties of algorithms (12) such as improved transient performance of the closed loop system and their ability to deal with nonconvexly parameterized models, these algorithms are not realizable in the form of differential equations, as they depend on unknown parameters explicitly. It was proposed in [30] to use special filters to estimate $\dot{\psi}$. While the approach of [30] is acceptable for systems with nonconvex parameterization, control system performance may be suboptimal due to estimation errors. The question is how to realize algorithms (12) in a form that depends neither on time-derivative $\dot{\psi}$ nor on its filtered estimate explicitly, nor on anything implying knowledge of unknown parameters $\boldsymbol{\theta}$. Our solution, as mentioned in Section 1, is to use the finite form (2) of adaptive algorithms instead of the differential form (12). In the next section we study under what conditions algorithms (12) can be represented in finite form (2).

3 Adaptive Algorithms in Finite Form

The outline of the section is as follows. We start from a general case and formulate the conditions ensuring the realization of algorithm (12) in finite form explicitly, i.e., without any filters and further transformations of the closed-loop system. The conditions we impose involve the existence of the solutions of a system of partial differential equations. It is nontrivial to check these assumptions for nonlinear model (3). That they hold, however, can be demonstrated for some special combinations of plant models and goal functions $\psi(\mathbf{x}, t)$. Further, we consider extension of the proposed method to a broader class of nonlinear systems including systems with low-triangular structure.

3.1 Explicit Realization

Let us assume that in addition to the Assumptions 1–5, that are sufficient for system (13) to have the properties P1)–P7), the following hold

Assumption 7 (Explicit realization condition) *For the given functions $\boldsymbol{\alpha}(\mathbf{x}, t)$ and $\psi(\mathbf{x}, t)$ there exists function $\Psi(\mathbf{x})$ such that the following hold:*

$$\Psi(\mathbf{x}) : \partial\Psi(\mathbf{x}, t)/\partial\mathbf{x}_2 = \psi(\mathbf{x}, t)(\partial\boldsymbol{\alpha}(\mathbf{x}, t)/\partial\mathbf{x}_2) \quad (16)$$

Then realizations of the adaptive scheme described by equations (12) follow from the next theorem.

Theorem 2 *Let Assumption 7 hold. Then there is a finite-form realization of the algorithms (12):*

$$\begin{aligned} \hat{\boldsymbol{\theta}}(\mathbf{x}, t) &= \Gamma(\hat{\boldsymbol{\theta}}_P(\mathbf{x}, t) + \hat{\boldsymbol{\theta}}_I(t)); \quad \hat{\boldsymbol{\theta}}_P(\mathbf{x}, t) = \psi(\mathbf{x}, t)\boldsymbol{\alpha}(\mathbf{x}, t) - \Psi(\mathbf{x}, t) \\ \dot{\hat{\boldsymbol{\theta}}}_I &= \varphi(\psi(\mathbf{x}, t))\boldsymbol{\alpha}(\mathbf{x}, t) + \partial\Psi(\mathbf{x}, t)/\partial t - \psi(\mathbf{x}, t)(\partial\boldsymbol{\alpha}(\mathbf{x}, t)/\partial t) - \\ &\quad (\psi(\mathbf{x}, t)L_{\mathbf{f}_1}\boldsymbol{\alpha}(\mathbf{x}, t) - L_{\mathbf{f}_1}\Psi(\mathbf{x}, t)) - (\psi(\mathbf{x}, t)L_{\mathbf{g}_1}\boldsymbol{\alpha}(\mathbf{x}, t) - L_{\mathbf{g}_1}\Psi(\mathbf{x}, t))u(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) \end{aligned} \quad (17)$$

Remark 1 It is easy to see from (17) and the theorem proof that realization of the algorithms

$$\dot{\boldsymbol{\theta}} = \Gamma(\dot{\psi} + \beta(\mathbf{x}, t))\boldsymbol{\alpha}(\mathbf{x}, t), \quad (18)$$

where $\beta(\mathbf{x}, t)$ is to guarantee at least the existence of solutions for the closed loop system, is also possible. Indeed, in order to realize these algorithms it is sufficient to replace equations for $\dot{\boldsymbol{\theta}}_I$ in (17) by the following:

$$\begin{aligned} \dot{\boldsymbol{\theta}}_I &= \beta(\mathbf{x}, t)\boldsymbol{\alpha}(\mathbf{x}, t) + \partial\Psi(\mathbf{x}, t)/\partial t - \psi(\mathbf{x}, t)(\partial\boldsymbol{\alpha}(\mathbf{x}, t)/\partial t) - \\ &\quad (\psi(\mathbf{x}, t)L_{\mathbf{f}_1}\boldsymbol{\alpha}(\mathbf{x}, t) - L_{\mathbf{f}_1}\Psi(\mathbf{x}, t)) - (\psi(\mathbf{x}, t)L_{\mathbf{g}_1}\boldsymbol{\alpha}(\mathbf{x}, t) - L_{\mathbf{g}_1}\Psi(\mathbf{x}, t))u(\mathbf{x}, \hat{\boldsymbol{\theta}}, t), \end{aligned} \quad (19)$$

One particular case of function $\beta(\mathbf{x}, t) = (1 + \delta(t))\psi(\mathbf{x}, t)$, $\delta : R_+ \rightarrow R_+$, $\delta \in C^0$ will be used later to show existence of the adaptive control algorithms for nonlinearly parameterized plants in the low-triangular form.

Theorem 2 provides us with an answer to the question of existence of realizable algorithms that satisfy differential equations (12), thus ensuring the properties formulated in Theorem 1 and Proposition 1. The disadvantage, however, is that the functions $\Psi(\mathbf{x}, t)$ in Assumption 7 are not easy to find. Existence of such functions itself is another nontrivial issue. For instance, if $\dim \mathbf{x}_2 = n$ and functions $\psi(\mathbf{x}, t)$, $\boldsymbol{\alpha}(\mathbf{x}, t)$ do not depend explicitly on time t , then the necessary conditions for the function $\Psi(\mathbf{x})$ to exist is the symmetry of all matrices $\frac{\partial}{\partial \mathbf{x}} \left(\psi(\mathbf{x}) \frac{\partial \alpha_i(\mathbf{x})}{\partial \mathbf{x}} \right)$, $i \in \{1, \dots, n\}$. Nevertheless, despite difficulties in finding those functions $\Psi(\mathbf{x}, t)$ that satisfy Assumption 7, there are several classes of dynamical systems with certain structural properties that immediately reduce Assumption 7 to more easily verifiable requirements.

Corollary 1 (Single-dimension uncertainty-dependent partition) *Let $\dim(\mathbf{x}_2) = 1$ and function $\psi(\mathbf{x}, t)\partial\boldsymbol{\alpha}(\mathbf{x}, t)/\partial x_n$ be Riemann-integrable with respect to x_n , i.e., the following integral exists*

$$\Psi(\mathbf{x}, t) = \int \psi(\mathbf{x}, t) \frac{\partial \boldsymbol{\alpha}(\mathbf{x}, t)}{\partial x_n} dx_n \quad (20)$$

Then there is a finite-form realization of algorithms (12).

Remark 2 Corollary 1 allows us to turn the problem of searching for a function $\Psi(\mathbf{x}, t)$ satisfying equation (16) into a problem of existence of the indefinite integral of a function with respect to a single scalar argument. It is clear from (20) that any one-dimensional system with integrable $\psi(x) \frac{\partial \alpha(x)}{\partial x}$ has a finite-form realization. An interesting example is the class of systems described by the following differential equations:

$$\begin{aligned} \dot{x}_i &= f_i(\mathbf{x}) + g_i(\mathbf{x})u, \quad i = 1, \dots, n-1 \\ \dot{x}_n &= f_n(\mathbf{x}) + \nu(\mathbf{x}, \boldsymbol{\theta}) + g_n(\mathbf{x})u, \end{aligned} \quad (21)$$

where function $\nu(\mathbf{x}, \boldsymbol{\theta})$ satisfies Assumption 5, which in turn is automatically satisfied if $\nu(\mathbf{x}, \boldsymbol{\theta})$ linearly parameterized or $\nu(\mathbf{x}, \boldsymbol{\theta}) = \nu(\mathbf{x}^T \boldsymbol{\theta})$ and $\nu(\cdot)$ is monotonic and belongs to a sector. In practice, the indefinite integral in (20) can also be replaced by $\Psi(\mathbf{x}, t) = \int_{x_n(0)}^{x_n(t)} \psi(\mathbf{x}, t) \frac{\partial \boldsymbol{\alpha}(\mathbf{x}, t)}{\partial x_n} dx_n$.

Equations of type (21) describe a class of dynamical systems in which the uncertainties are concentrated in a single equation. There are many mechanical systems described by equations (21) satisfying Assumption

7 (e.g., the simple classical equations for shaft dynamics with unknown load torque, rotating platforms and pendulums). For the case where the uncertainty is a single scalar and function $\alpha(\mathbf{x})$ can be chosen as the goal function $\psi(\mathbf{x}) = \alpha(\mathbf{x})$, finite-form realization is also possible ($\Psi(\mathbf{x}) = \frac{1}{2}\alpha^2(\mathbf{x})$). Another class of dynamical systems that automatically satisfy Assumption 7 is given by the following corollary.

Corollary 2 (Independence on partition \mathcal{L}_2) *Let function $\alpha(\mathbf{x}, t)$ be independent on \mathcal{L}_2 , i. e., for any $\mathbf{x}_2 \in \mathcal{L}_2$ the following holds: $\partial\alpha(\mathbf{x}, t)/\partial\mathbf{x}_2 = \partial\alpha(\mathbf{x}_1 \oplus \mathbf{x}_2, t)/\partial\mathbf{x}_2 = 0$ then there is a finite-form realization of algorithms (12).*

Corollary 2 conditions are equivalent to the fact that the plant dynamics can be described by system

$$\begin{aligned}\dot{\mathbf{x}}_1 &= \mathbf{f}_1(\mathbf{x}_1 \oplus \mathbf{x}_2) + \mathbf{g}_1(\mathbf{x}_1 \oplus \mathbf{x}_2)u \\ \dot{\mathbf{x}}_2 &= \mathbf{f}_2(\mathbf{x}_1 \oplus \mathbf{x}_2) + \boldsymbol{\nu}(\mathbf{x}_1, \boldsymbol{\theta}) + \mathbf{g}_2(\mathbf{x}_1 \oplus \mathbf{x}_2)u\end{aligned}\tag{22}$$

and that $\partial\psi(\mathbf{x}_1 \oplus \mathbf{x}_2, t)/\partial\mathbf{x}_2 = \lambda(\mathbf{x}_1, t)$, where $\lambda : R^m \times R \rightarrow R^{n-m}$ is a differentiable function with known derivative $\partial\lambda(\mathbf{x}_1, t)/\partial t$. Therefore it is possible to derive from Corollary 2 that every error model: $\dot{\psi} = -\varphi(\psi) + z(\boldsymbol{\omega}(t), \boldsymbol{\theta}) - z(\boldsymbol{\omega}(t), \hat{\boldsymbol{\theta}})$, where $\boldsymbol{\omega}(t) : R \rightarrow R^n$, $\boldsymbol{\omega} \in C^1$ is a function with known time-derivatives $\dot{\boldsymbol{\omega}}(t)$, satisfies the sufficient conditions for realization of algorithm (12) in finite form. Indeed, this follows directly from Assumption 5, as functions $\alpha(\mathbf{x}, t)$ in this case are independent of \mathbf{x} . Therefore if the derivatives $\dot{\alpha}(t)$ are known, the finite form realization follows immediately from

$$\begin{aligned}\hat{\boldsymbol{\theta}}(\mathbf{x}, t) &= \Gamma(\hat{\boldsymbol{\theta}}_P(\mathbf{x}, t) + \hat{\boldsymbol{\theta}}_I); \quad \hat{\boldsymbol{\theta}}_P(\mathbf{x}, t) = \psi(\mathbf{x}, t)\boldsymbol{\alpha}(t) \\ \dot{\hat{\boldsymbol{\theta}}}_I &= \varphi(\psi(\mathbf{x}, t))\boldsymbol{\alpha}(\mathbf{x}, t) - \psi(\mathbf{x}, t)\dot{\boldsymbol{\alpha}}(t)\end{aligned}$$

This fact, along with decomposition (22), will be used later in Section 3.2.

So far, simplified conditions for the existence of the adaptive algorithms in finite form were derived from Theorem 2 for those classes of nonlinear systems that have certain structural properties, such as single dimension uncertainty-dependent partition (Corollary 1 and equation (21)) or independence of $z(\mathbf{x}, \boldsymbol{\theta}, t)$ on uncertainty-dependent partition \mathbf{x}_2 (Corollary 2). These structural properties allowed us to reduce Assumption 7 to integrability of a function with respect to a single scalar argument for a class of nonlinear systems. Taking these results into account, in the next section we present a technique that allows us to extend our method to a broader class of systems.

3.2 Asymptotic Design via Embedding

The main idea behind the extension of our results to a broader class of nonlinear systems is as follows. Instead of trying to find a general solution of equation (16) in Assumption 7 (which is a nontrivial task even if such solution exists), we transform the original equations into a form that satisfies much weaker requirements considered in Corollaries 1 and 2. This transformation should not necessarily be a one-to-one diffeomorphism, but the control goal reaching in the new state space should guarantee reaching the control goal of the original system. One

way to assure this is to embed the original system dynamics into one of a higher order, for which a finite form realization of the adaptive control algorithms is possible.

Let us represent the partitioned system (4) in the following way:

$$\begin{aligned}\dot{\mathbf{x}}_1 &= \mathbf{f}_1(\mathbf{x}) + \mathbf{g}_1(\mathbf{x})u \\ \dot{\mathbf{x}}'_2 &= \mathbf{f}'_2(\mathbf{x}) + \boldsymbol{\nu}'(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{g}'_2(\mathbf{x})u; \quad \dot{\mathbf{x}}''_2 = \mathbf{f}''_2(\mathbf{x}) + \boldsymbol{\nu}''(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{g}''_2(\mathbf{x})u,\end{aligned}\tag{23}$$

where $\mathbf{x}'_2 \oplus \mathbf{x}''_2 = \mathbf{x}_2$, $\dim \mathbf{x}'_2 = m_1$, $\dim \mathbf{x}''_2 = n - m - m_1$, $0 \leq m_1 \leq n - m$. Using the notations above, we introduce the following assumption

Assumption 8 *There exist*

- 1) a partition of the state vector \mathbf{x} : $\mathbf{x} = \mathbf{x}_1 \oplus \mathbf{x}'_2 \oplus \mathbf{x}''_2$,
- 2) a system of differential equations

$$\dot{\boldsymbol{\xi}} = \mathbf{f}_\xi(\mathbf{x}, \boldsymbol{\xi}, t); \quad \mathbf{y}_\xi = \mathbf{h}_\xi(\boldsymbol{\xi}),\tag{24}$$

$$\boldsymbol{\xi} \in R^r, \quad \mathbf{f}_\xi : R^n \times R^r \times R_+ \rightarrow R^r, \quad \mathbf{f}_\xi \in C^1; \quad \mathbf{h}_\xi : R^r \rightarrow R^{n-m-m_1}, \quad \mathbf{h}_\xi \in C^1;$$

- 3) a function $\Psi(\tilde{\mathbf{x}}, t) \in C^1$, $\tilde{\mathbf{x}} = \mathbf{x}_1 \oplus \mathbf{x}'_2 \oplus \mathbf{h}_\xi$ such that the following conditions hold

$$\partial \Psi(\tilde{\mathbf{x}}, t) / \partial \mathbf{x}'_2 = \psi(\tilde{\mathbf{x}}, t) (\partial \boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) / \partial \mathbf{x}'_2)\tag{25}$$

$$\mathbf{x} \in L_\infty \Rightarrow \boldsymbol{\xi} \in L_\infty\tag{26}$$

for any $\boldsymbol{\theta} \in \Omega_\theta$ and $t \in R_+$ along the solutions of the original system (3).

In addition to Assumption 8, we would like to formulate two alternative assumptions which, if satisfied, will result in two different adaptation schemes with different performance and robustness properties.

Assumption 9 *Let system (24) be given and*

$$z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t) \in L_\infty; \quad (\psi(\mathbf{x}, t) - \psi(\tilde{\mathbf{x}}, t)) (\partial \boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) / \partial \mathbf{x}'_2) \boldsymbol{\nu}'(\mathbf{x}, \boldsymbol{\theta}) \in L_\infty.\tag{27}$$

along the solutions of (3), (24).

Assumption 10 *Let system (24) be given and $\partial \boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) / \partial \mathbf{x}'_2 \equiv 0$, or $\psi(\mathbf{x}, t) = \psi(\tilde{\mathbf{x}}, t)$. Furthermore, let $z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t) \in L_2$ for any $\boldsymbol{\theta} \in \Omega_\theta$, $t > 0$ along the solutions of (3), (24).*

Sufficient conditions for the desired embedding follow from the next theorem.

Theorem 3 (Embedding Theorem) *Let function $\psi(\mathbf{x}, t)$ be given and Assumptions 1–6, 8 hold for system (3). Then for the extended system*

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) + \boldsymbol{\vartheta}(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{g}(\mathbf{x})u \\ \dot{\boldsymbol{\xi}} &= \mathbf{f}_\xi(\mathbf{x}, \boldsymbol{\xi}, t); \quad \mathbf{y}_\xi = \mathbf{h}_\xi(\boldsymbol{\xi}),\end{aligned}\tag{28}$$

there exists control function $u(\mathbf{x}, \mathbf{h}_\xi, \hat{\boldsymbol{\theta}}, t)$

$$u(\mathbf{x}, \mathbf{h}_\xi, \boldsymbol{\theta}, t) = (L_{\mathbf{g}(\mathbf{x})}\psi(\mathbf{x}, t))^{-1} \left(-\varphi(\psi) - L_{\mathbf{f}}\psi(\mathbf{x}, t) - L_{\boldsymbol{\theta}(\tilde{\mathbf{x}}, \hat{\boldsymbol{\theta}})}\psi(\tilde{\mathbf{x}}, t) - \partial\psi(\mathbf{x}, t)/\partial t \right) \quad (29)$$

and adaptation algorithms⁴: $\hat{\boldsymbol{\theta}}(\tilde{\mathbf{x}}, t) = \Gamma(\hat{\boldsymbol{\theta}}_P(\tilde{\mathbf{x}}, t) + \hat{\boldsymbol{\theta}}_I(t))$, $\Gamma > 0$ such that the following statements hold:

P8) if $|\varphi(\psi)| \geq K|\psi|$, $K > 0$ and Assumption 9 holds then $\psi(\mathbf{x}, t), \mathbf{x}, \boldsymbol{\xi}, \hat{\boldsymbol{\theta}} \in L_\infty$;

P9) if Assumption 10 holds then $\psi(\mathbf{x}, t) \in L_2 \cap L_\infty, \dot{\psi} \in L_2$, $z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \hat{\boldsymbol{\theta}}, t) \in L_2$, $\mathbf{x}, \boldsymbol{\xi} \in L_\infty$;

if in addition derivatives $\partial\psi(\mathbf{x}, t)/\partial\mathbf{x}$, $\partial\psi(\mathbf{x}, t)/\partial t$ are uniformly bounded in t and $z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t) \in L_\infty$ then $\dot{\psi} \in L_\infty$, $z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \hat{\boldsymbol{\theta}}, t) \in L_\infty$, $\lim_{t \rightarrow \infty} \psi(\mathbf{x}(t), t) = 0$.

Theorem 3 states not only existence of the adaptive control algorithms but also provides us with exact equations for the adaptive control function. These equations are given by (29), (77), which guarantee P8), and (29), (81) or (82) ensuring P9) for $\partial\boldsymbol{\alpha}(\tilde{\mathbf{x}})/\partial\mathbf{x}'_2 \equiv 0$ or $\psi(\tilde{\mathbf{x}}) = \psi(\mathbf{x})$ respectively.

Although Theorem 3 guarantees reaching of the control goal and ascertains performance improvement (property P9)), it does not ensure the same properties of adaptive control algorithms as Theorem 2 does. On the other hand, the ability to deal with nonconvex parameterized systems is preserved, except for cases that do not satisfy Assumption 6. The drawbacks of this narrower class of nonlinearly parameterized functions in the plant right-hand side and a slight degradation in performance are compensated by relaxing the requirement (16) of Assumption 7. Notice also that the difference in guaranteed performance reflected in P8) and P9) has the consequence that the dimensions of vectors \mathbf{h}_ξ are likely to be different in the both cases. Indeed, to ensure equality $\partial\boldsymbol{\alpha}(\tilde{\mathbf{x}}, t)/\partial\mathbf{x}'_2 \equiv 0$ for arbitrary smooth function $\boldsymbol{\alpha}(\cdot)$, we must replace the whole vector \mathbf{x}_2 by $\mathbf{h}_\xi(\boldsymbol{\xi})$. Therefore, in principle, embedding of the original system dynamics into one of a higher order is desired if improved performance and extended applicability are required.

Theorem 3 offers a possible way to facilitate the search for function $\Psi(\mathbf{x}, t)$ satisfying partial differential equation (16) as defined in Assumption 7. We replace the problem by one of searching for the embedding (28) which satisfies Assumption 8 and 9 or 10. The main obstacle, finding a solution to equation (16), is replaced with problem (25), the complexity⁵ of which should be reduced, as $\dim \mathbf{x}'_2 < \dim \mathbf{x}_2$ if embedding into the higher-order dynamics is used.

Indeed, according to Assumption 8 and notations introduced above, the dynamics of the extended system can be described as

$$\begin{aligned} \dot{\mathbf{x}}_1 &= \mathbf{f}_1(\mathbf{x}) + \mathbf{g}_1(\mathbf{x})u, & \dot{\mathbf{x}}_{2'} &= \mathbf{f}_{2'}(\mathbf{x}) + \boldsymbol{\nu}'(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{g}_{2'}(\mathbf{x})u \\ \dot{\mathbf{h}}_\xi &= \frac{\partial \mathbf{h}_\xi}{\partial \boldsymbol{\xi}} \mathbf{f}_\xi(\mathbf{x}, \boldsymbol{\xi}, t), & \dot{\mathbf{x}}_{2''} &= \mathbf{f}_{2''}(\mathbf{x}) + \boldsymbol{\nu}''(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{g}_{2''}(\mathbf{x})u, \end{aligned} \quad (30)$$

where vector $\mathbf{x}_1 \oplus \mathbf{h}_\xi$ stands for the uncertainty-independent partition in the extended state space, and vector \mathbf{x}'_2 is chosen to satisfy equation (25). Observe that function $z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t)$ is independent of \mathbf{x}''_2 and $\dim \mathbf{h}_\xi = \dim \mathbf{x}''_2$.

Then for any \mathbf{h}_ξ : $\dim \mathbf{h}_\xi > 0$, we can conclude that $\dim \mathbf{x}'_2 < \dim \mathbf{x}_2 = \dim \mathbf{x}'_2 \oplus \mathbf{x}''_2$.

⁴The adaptation algorithms that guarantee properties P8) and P9) are given by equations (77) and (81), (82) respectively in Appendix 2.

⁵Here reduced complexity means that the number of equations in the system is reduced.

Notice also that, by the appropriate choice of the dimensions of vectors ξ and \mathbf{h}_ξ ($\dim \mathbf{h}_\xi = \dim \mathbf{x}_2''$) in (24), the dimension of vector \mathbf{x}_2' can be reduced to unity. Alternatively, we may try to annihilate the partial derivative $\frac{\partial \alpha(\tilde{\mathbf{x}}, t)}{\partial \mathbf{x}_2'}$ in (25). Hence, eventually either Corollary 1 or Corollary 2 conditions will be satisfied for the extended system (30). This, in turn, implies that we can replace assumption (25) by a weaker requirement, such as integrability of the function with respect to a single scalar argument.

After obtaining computable function $\Psi(\tilde{\mathbf{x}}, t)$, the remaining problem is that we should be able to find an extension (24) that guarantees properties (27) and (26) for the given partition $\tilde{\mathbf{x}} = \mathbf{x}_1 \oplus \mathbf{x}_2' \oplus \mathbf{h}_\xi$. If such an extension exists, then Assumption 8 is automatically satisfied, and adaptive control algorithms follow immediately from Theorem 3.

Finding extension (24) that ensures boundedness (and square integrability) of the differences $z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t)$, $(\psi(\mathbf{x}, t) - \psi(\tilde{\mathbf{x}}, t))(\partial \alpha(\tilde{\mathbf{x}}, t)/\partial \mathbf{x}_2')\boldsymbol{\nu}'(\mathbf{x}, \boldsymbol{\theta})$ is not an easy problem – taking into account that partition \mathbf{x}_2'' is also uncertainty-dependent. It is possible to solve it using specially designed *adaptive* or *high-gain* auxiliary subsystems that track the reference signals \mathbf{x}_2'' with the desired performance: $z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t), (\psi(\mathbf{x}, t) - \psi(\tilde{\mathbf{x}}, t))\frac{\partial \alpha(\tilde{\mathbf{x}}, t)}{\partial \mathbf{x}_2'}\boldsymbol{\nu}'(\mathbf{x}, \boldsymbol{\theta}) \in L_2 \cap L_\infty$. If, for example, partition \mathbf{x}_2'' is linearly parameterized (i.e., $\boldsymbol{\nu}''(\mathbf{x}, \boldsymbol{\theta}) = \boldsymbol{\eta}''(\mathbf{x})\boldsymbol{\theta}$), functions $z(\mathbf{x}, \boldsymbol{\theta}, t), \psi(\mathbf{x}, t)\frac{\partial \alpha(\tilde{\mathbf{x}}, t)}{\partial \mathbf{x}_2'}\boldsymbol{\nu}'(\mathbf{x}, \boldsymbol{\theta})$ are locally Lipschitz in \mathbf{x}_2'' and for any $\boldsymbol{\theta} \in \Omega_\theta$ the following inequalities hold: $|z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t)| \leq \lambda_1(\mathbf{x}, \xi, t)\|\mathbf{x}_2'' - \mathbf{h}_\xi(\xi)\|$, $\|(\psi(\mathbf{x}, t) - \psi(\tilde{\mathbf{x}}, t))\frac{\partial \alpha(\tilde{\mathbf{x}}, t)}{\partial \mathbf{x}_2'}\boldsymbol{\nu}'(\mathbf{x}, \boldsymbol{\theta})\| \leq \lambda_2(\mathbf{x}, \xi, t)\|\mathbf{x}_2'' - \mathbf{h}_\xi(\xi)\|$ then the suitable extension is defined by the following system:

$$\begin{aligned}\dot{\xi}_1 &= \mathbf{f}_2''(\mathbf{x}) + \boldsymbol{\eta}''(\mathbf{x})\xi_2 + \bar{\lambda}(\mathbf{x}, \xi, t)(\mathbf{x}_2'' - \xi_1) + \mathbf{g}_2''(\mathbf{x})u \\ \dot{\xi}_2 &= \Gamma_1(\mathbf{x}_2'' - \xi_1)^T \boldsymbol{\eta}''(\mathbf{x}), \quad \Gamma_1 > 0, \quad \mathbf{h}_\xi(\xi) = \xi_1,\end{aligned}\tag{31}$$

where $\xi = \xi_1 \oplus \xi_2$ and $\bar{\lambda}(\mathbf{x}, \xi, t) = \lambda_1^2(\mathbf{x}, \xi, t) + \lambda_2^2(\mathbf{x}, \xi, t)$. To show this, it is sufficient to consider the following Lyapunov's candidate: $V(\mathbf{x}, \xi) = 0.5\|(\mathbf{x}_2'' - \xi_1)\|^2 + 0.5\|\boldsymbol{\theta} - \xi_2\|_{\Gamma^{-1}}^2$ and observe that $\dot{V} \leq -\bar{\lambda}(\mathbf{x}, \xi, t)\|\mathbf{x}_2'' - \xi_1\|^2 \leq -(z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t))^2 \leq 0$.

For a class of nonlinear systems with low-triangular structure

$$\begin{aligned}\dot{x}_i &= f_i(x_1, \dots, x_i, \boldsymbol{\theta}_i) + x_{i+1}, \quad i = 1, \dots, n-1, \\ \dot{x}_n &= f_n(x_1, \dots, x_n, \boldsymbol{\theta}_n) + u + \varepsilon(t), \quad \varepsilon(t) \in L_2, \quad \boldsymbol{\theta}_i \in \Omega_\theta\end{aligned}\tag{32}$$

the suitable extension is guaranteed by Lemma 3 in Appendix 1. Then, combining the results formulated in Theorem 3 and Lemma 3, it is possible to show that our approach can be extended to a broad class of systems like those given by equations (32). Let functions $f_i(\cdot)$ in (32) satisfy the following assumption

Assumption 11 *Let there exist smooth functions $\bar{D}_i(\cdot) : R^i \times R^i \times R \rightarrow R$ such that for any $\boldsymbol{\theta}_i \in \Omega_\theta$ the following holds: $(f_i(x_1, \dots, x_i, \boldsymbol{\theta}_i) - f_i(x'_1, \dots, x'_i, \boldsymbol{\theta}_i))^2 \leq \bar{D}_i^2(\mathbf{x}_i, \mathbf{x}'_i)\|\mathbf{x}_i - \mathbf{x}'_i\|^2$, $\mathbf{x}_i = (x_1, \dots, x_i)^T$, $\mathbf{x}'_i = (x'_1, \dots, x'_i)^T$.*

It is clear that Assumption 11 holds for those functions $f_i(x_1, \dots, x_i, \boldsymbol{\theta}_i)$ that are, for example, Lipschitz in \mathbf{x} . The results for low-triangular systems (32) are formulated in the next theorem

Theorem 4 (Finite Forms for Low-Triangular Systems) *Let system (32) and goal function $\psi(x_1) = 0$ be given, and there exist functions $\alpha_i(x_1, \dots, x_i)$ such that Assumptions 5, 6 hold for the functions $f_i(x_1, \dots, x_i, \theta_i)$ in (32) respectively. Furthermore, let $f_i(x_1, \dots, x_i, \theta_i)$ satisfy Assumption 11, $\alpha_i(x_1, \dots, x_i)$, $f_i(x_1, \dots, x_i, \theta_i)$, $i = 1, \dots, n$, $\psi_1(x_1)$ be smooth and the following condition holds: $\psi(x_1) \in L_\infty \Rightarrow x_1 \in L_\infty$.*

Then there exist an auxiliary system

$$\dot{\xi} = f_\xi(\mathbf{x}, \xi, \nu), \quad \xi_0 \in R^n, \quad \dot{\nu} = f_\nu(\mathbf{x}, \xi, \nu), \quad \nu_0 \in R^m, \quad (33)$$

as well as smooth functions $\psi_i(x_i, t)$, $i = 1, \dots, n$, $\hat{\theta}_P(\mathbf{x}, \xi)$, control $u(\mathbf{x}, \hat{\theta}, \xi, \nu)$, and adaptation algorithm

$$\hat{\theta}(\mathbf{x}, \xi, \hat{\theta}_I) = \gamma(\hat{\theta}_P(\mathbf{x}, \xi) + \hat{\theta}_I), \quad \gamma > 0, \quad \dot{\hat{\theta}}_I = f_{\hat{\theta}}(\mathbf{x}, \hat{\theta}, \xi, \nu),$$

such that

- 1) $\psi_i(x_i, t), \psi \in L_2 \cap L_\infty, \dot{\psi}, \dot{\psi}_i \in L_2, i = 1, \dots, n$
- 2) $\hat{\theta} \in L_\infty$ and $u(\mathbf{x}, \hat{\theta}, \xi, \nu) - u(\mathbf{x}, \theta_n, \xi, \nu) \in L_2$
- 3) $\mathbf{x}, \xi, \nu \in L_\infty$
- 4) if $\varepsilon(t) \in L_\infty$ then $\dot{\psi}, \dot{\psi}_i \in L_\infty$, and $\lim_{t \rightarrow \infty} \psi(x_1(t)) = 0, \quad \lim_{t \rightarrow \infty} \psi_i(x_i(t), t) = 0, i = 1, \dots, n$.

Theorem 4 extends the applicability of algorithms in finite form to systems described by equation (32). Relying entirely on Lemma 3 (Appendix 1) and Theorem 3, Theorem 4 allows us to design adaptive control algorithms for cascades with nonlinear parameterization without the need for damping nonlinearities. However, performance is weaker. For instance, decrease (non-increase) of the term $\|\theta - \hat{\theta}(t)\|_{\Gamma^{-1}}^2$ is not guaranteed in this case. Nevertheless, adaptive control algorithms in finite form, in addition to their ability to deal with nonlinear parameterization, still guarantee certain improvements in performance. For instance, square integrability of the control effort due to adaptation (statement 2) of the theorem) and $\psi_i(x_i, t), \dot{\psi}_i \in L_2 \cap L_\infty$ are ensured. In the next section we illustrate our method with the examples.

4 Examples

Let us consider the following system:

$$\dot{x}_1 = x_1^2 \theta_0 + x_2; \quad \dot{x}_2 = x_1 \theta_1 + x_2 \theta_2 + u, \quad (34)$$

where parameters θ_0, θ_1 and θ_2 are assumed to be unknown. The control goal is to steer the system towards the following manifold: $x_1 - 1 = 0$. To design adaptive algorithms in finite form for system (34), we follow the steps of Theorem 4 proof:

- 1) *Intermediate control design.* Derive control function $u_1(x_1, \hat{\theta}_0)$ such that for the reduced system

$$\dot{x}_1 = x_1^2 \theta_0 + u_1(x_1, \hat{\theta}_0) + \varepsilon_1(t), \quad \varepsilon_1(t) \in L_2; \quad \hat{\theta}_0 = \hat{\theta}_{0,P}(x_1) + \hat{\theta}_{0,I}(t)$$

reaching of the control goal is guaranteed: $\psi(x_1(t)) = x_1(t) - 1 \rightarrow 0$ as $t \rightarrow \infty$. Moreover, function $u_1(x_1, \hat{\theta}_0(x_1, \hat{\theta}_{0,I}))$ should ensure that $\psi, \dot{\psi} \in L_2$.

2) *Embedding.* Extend the system dynamics (or embed it into) with auxiliary system

$$\dot{\xi} = f_{\xi}(\mathbf{x}, \xi, \nu); \quad \dot{\nu} = f_{\nu}(\mathbf{x}, \xi, \nu) \quad (35)$$

in order to guarantee that

$$u(x_1, \hat{\theta}_0(x_1, \hat{\theta}_{0,I})) - u(\xi, \hat{\theta}_0(\xi, \hat{\theta}_{0,I})) \in L_2, \quad x_1 - \xi \in L_2 \quad (36)$$

3) *Control function design.* Introduce new goal function $\psi_2(x_2, t) = x_2 - u_1(\xi, \hat{\theta}_0(\xi, \hat{\theta}_{0,I}))$ and derive control function $u(x_1, x_2, \xi, t)$ such that $\dot{\psi}_2 \in L_2$, $\psi_2 \in L_2 \cap L_{\infty}$. The last automatically implies that $\dot{x}_1 = x_1^2 \theta_0 + x_2 = x_1^2 \theta_0 + u_1(x_1, \hat{\theta}_0(x_1, \hat{\theta}_{0,I})) + \mu(t)$, where $\mu(t) = x_2 - u_1(x_1, \hat{\theta}_0(x_1, \hat{\theta}_{0,I})) = (x_2 - u_1(\xi, \hat{\theta}_0(\xi, \hat{\theta}_{0,I}))) + (u_1(\xi, \hat{\theta}_0(\xi, \hat{\theta}_{0,I})) - u_1(x_1, \hat{\theta}_0(x_1, \hat{\theta}_{0,I}))) \in L_2$. Therefore, according to the choice of function $u_1(x_1, \hat{\theta}_0(x_1, \hat{\theta}_{0,I}))$, control $u(x_1, x_2, \xi, t)$ guarantees that $\psi(x_1(t)) \rightarrow 0$ as $t \rightarrow \infty$, $\psi, \dot{\psi} \in L_2$.

We begin by determining the function $u(x_1, \hat{\theta}_0(x_1, \hat{\theta}_{0,I}))$. Let $u_1(x_1, \hat{\theta}_0) = -(x_1 - 1) - \hat{\theta}_0 x_1^2$, where $\hat{\theta}_0$ satisfies the following differential equation:

$$\dot{\hat{\theta}}_0 = (x_1 - 1 + \dot{x}_1)x_1^2 \quad (37)$$

It follows from Lemma 2 that control function $u_1(x_1, \hat{\theta}_0)$ with algorithm (37) guarantee that $\psi, \dot{\psi} \in L_2$, $\psi(x_1(t)) \rightarrow 0$ as $t \rightarrow \infty$. According to Theorem 2, finite form realization of (37) can be given as follows: $\hat{\theta}_0(x_1, \hat{\theta}_{0,I}(t)) = 1/3x_1^3 + \hat{\theta}_{0,I}(t)$; $\dot{\hat{\theta}}_{0,I} = (x_1 - 1)x_1^2$. Substituting this into $u_1(x_1, \hat{\theta}_0)$ we get the following expression for $u_1(\cdot)$:

$$u_1(x_1, \hat{\theta}_0(x_1, \hat{\theta}_{0,I})) = -(x_1 - 1) - 1/3x_1^5 - x_1^2 \hat{\theta}_{0,I}(t); \quad \dot{\hat{\theta}}_{0,I} = \psi(x_1)\alpha_1(x_1) = (x_1 - 1)x_1^2. \quad (38)$$

Thus step 1 is completed.

Let us design system (35) which guarantees that (36) holds for function (38). First consider the difference:

$$u(x_1, \hat{\theta}_0(x_1, \hat{\theta}_{0,I})) - u(\xi, \hat{\theta}_0(\xi, \hat{\theta}_{0,I})) = -(x_1 - \xi)(1 + (x_1 + \xi)\hat{\theta}_{0,I} + 1/3(x_1^4 + x_1^3\xi + x_1^2\xi^2 + x_1\xi^3 + \xi^4)) \quad (39)$$

and denote $F(x_1, \xi, \hat{\theta}_{0,I}) = (1 + (x_1 + \xi)\hat{\theta}_{0,I} + 1/3(x_1^4 + x_1^3\xi + x_1^2\xi^2 + x_1\xi^3 + \xi^4))$. It follows from Lemma 3 that there exists system (35) such that condition (36) holds. In fact, this system can be given by the following equation

$$\dot{\xi} = (x_1 - \xi)(F^2(x_1, \xi, \hat{\theta}_{0,I}) + 1) + x_1^2 \hat{\theta}_{\xi} + x_2 \quad (40)$$

where $\hat{\theta}_{\xi}$ satisfies the following differential equation $\dot{\hat{\theta}}_{\xi} = (x_1 - \xi + \dot{x}_1 - \dot{\xi})x_1^2$. Finite form realization of this algorithm⁶ follows from Theorem 2, and it can be written as:

$$\hat{\theta}_{\xi} = 1/3x_1^3 + \hat{\theta}_{\xi,I}; \quad \dot{\hat{\theta}}_{\xi,I} = (x_1 - \xi)x_1^2 - x_1^2((x_1 - \xi)(F^2(x_1, \xi, \hat{\theta}_{0,I}) + 1) + x_1^2 \hat{\theta}_{\xi} + x_2) \quad (41)$$

⁶Introduction of algorithms (41) is not necessary here because the original system is linearly parameterized, and condition (36) can be satisfied even with conventional (gradient) adaptation schemes. Nevertheless, we would like to keep the consistency of our current calculations with those steps made in the proof of Theorem 4 in order to illustrate what would happen if the right hand sides are nonlinearly parameterized.

Taking into account (41) and (40) system (35) which ensures (36) can be represented as follows

$$\begin{aligned}\dot{\xi} &= (x_1 - \xi)(F^2(x_1, \xi, \hat{\theta}_{0,I}) + 1) + \frac{1}{3}x_1^5 + \hat{\theta}_{\xi,I}(t)x_1^2 + x_2 \\ \dot{\hat{\theta}}_{\xi,I} &= (x_1 - \xi)x_1^2 - x_1^2((x_1 - \xi)(F^2(x_1, \xi, \hat{\theta}_{0,I}) + 1) + \frac{1}{3}x_1^5 + \hat{\theta}_{\xi,I}(t)x_1^2 + x_2).\end{aligned}\quad (42)$$

Therefore, step 2 is completed as well. To conclude the controller design let us consider new target manifold $x_2 - u_1(\xi, \hat{\theta}_0(\xi, \hat{\theta}_{0,I})) = 0$ and goal function $\psi_2(x_2, t) = x_2 - u_1(\xi, \hat{\theta}_0(\xi, \hat{\theta}_{0,I})) = x_2 + \xi - 1 + \frac{1}{3}\xi^5 + \hat{\theta}_{0,I}\xi^2$. Let us write function $\psi_2(\cdot)$ derivative with respect to time t :

$$\begin{aligned}\dot{\psi}_2 &= \dot{x}_2 - \frac{\partial u_1(\xi, \hat{\theta}_0(\xi, \hat{\theta}_{0,I}))}{\partial \xi} \dot{\xi} - \frac{\partial u_1(\xi, \hat{\theta}_0(\xi, \hat{\theta}_{0,I}))}{\partial \hat{\theta}_{0,I}} \dot{\hat{\theta}}_{0,I} = x_1\theta_1 + x_2\theta_2 + u + \xi^2(x_1 - 1)x_1^2 + \\ &\quad (1 + \frac{5}{3}\xi^4 + 2\xi\hat{\theta}_{0,I})((x_1 - \xi)(F^2(x_1, \xi, \hat{\theta}_{0,I}) + 1) + \frac{1}{3}x_1^5 + \hat{\theta}_{\xi,I}(t)x_1^2 + x_2)\end{aligned}$$

Therefore, control function

$$\begin{aligned}u &= -\xi\hat{\theta}_1 - x_2\hat{\theta}_2 - \xi^2(x_1 - 1)x_1^2 - (x_2 + \xi - 1 + \frac{1}{3}\xi^5 + \hat{\theta}_{0,I}\xi^2) - \\ &\quad (1 + \frac{5}{3}\xi^4 + 2\xi\hat{\theta}_{0,I})((x_1 - \xi)(F^2(x_1, \xi, \hat{\theta}_{0,I}) + 1) + \frac{1}{3}x_1^5 + \hat{\theta}_{\xi,I}(t)x_1^2 + x_2)\end{aligned}\quad (43)$$

results in the following error model: $\dot{\psi}_2 = -\psi_2(x_2, t) + x_1\theta_1 + x_2\theta_2 - x_1\hat{\theta}_1 - x_2\hat{\theta}_2$. Taking into account condition (36), we can rewrite derivative $\dot{\psi}_2$ as $\dot{\psi}_2 = -\psi_2(x_2, t) + \xi\theta_1 + x_2\theta_2 - \xi\hat{\theta}_1 - x_2\hat{\theta}_2 + \varepsilon(t)$, where $\varepsilon(t) = (x_1 - \xi)\theta_1 \in L_2$. It follows from Lemma 2 that adaptation algorithm

$$\dot{\hat{\theta}}_1 = (\psi_2(x_2, t) + \dot{\psi}_2)\alpha_1(\xi); \quad \dot{\hat{\theta}}_2 = (\psi_2(x_2, t) + \dot{\psi}_2)\alpha_2(x_2), \quad \alpha_1(\xi) = \xi, \quad \alpha_2(x_2) = x_2$$

guarantees $\psi_2 \in L_2 \cap L_\infty$ and $\dot{\psi}_2 \in L_2$. Realization of algorithms (44) can be obtained from Theorem 2:

$$\begin{aligned}\hat{\theta}_1(x_2, \xi, \hat{\theta}_{0,I}, t) &= (x_2 + \xi - 1 + \frac{1}{3}\xi^5 + \hat{\theta}_{0,I}\xi^2)\xi + \hat{\theta}_{1,I}(t); \quad \dot{\hat{\theta}}_{1,I} = (x_2 + \xi - 1 + \frac{1}{3}\xi^5 + \hat{\theta}_{0,I}\xi^2)(\xi - \dot{\xi}) \\ \hat{\theta}_2(x_2, \xi, \hat{\theta}_{0,I}, t) &= \frac{x_2^2}{2} + \hat{\theta}_{2,I}(t); \quad \dot{\hat{\theta}}_{2,I} = (x_2 + \xi - 1 + \frac{1}{3}\xi^5 + \hat{\theta}_{0,I}\xi^2)x_2 + \frac{\partial \Psi_2}{\partial \xi} \dot{\xi} + \frac{\partial \Psi_2}{\partial \hat{\theta}_{0,I}} \dot{\hat{\theta}}_{0,I},\end{aligned}\quad (44)$$

where $\Psi_2(x_2, \xi, \hat{\theta}_{0,I}) = \int \psi_2(x_2, t) \frac{\partial \alpha_2(x_2)}{\partial x_2} dx_2 = \frac{x_2^2}{2} + (\xi - 1 + \frac{1}{3}\xi^5 + \hat{\theta}_{0,I}\xi^2)x_2$.

We would also like to compare performance of the proposed adaptation scheme with adaptive backstepping control algorithms. Adaptive backstepping design for system (34) according to [8] results in control algorithm:

$$\begin{aligned}u_1 &= -2x_2 - (x_1 - 1) - \hat{\theta}_3x_1^2 - x_1^4(x_1 - 1) - 2\hat{\theta}_3x_1x_2 - (x_1^2 + 2\hat{\theta}_3x_1^3)\hat{\theta} - x_1\hat{\theta}_1 - x_2\hat{\theta}_2 \\ \dot{\hat{\theta}} &= (x_2 + x_1 - 1 + \hat{\theta}_3x_1^2)x_1^2(1 + 2\hat{\theta}_3x_1); \quad \dot{\hat{\theta}}_1 = (x_2 + x_1 - 1 + \hat{\theta}_3x_1^2)x_1 \\ \dot{\hat{\theta}}_2 &= (x_2 + x_1 - 1 + \hat{\theta}_3x_1^2)x_2; \quad \dot{\hat{\theta}}_3 = (x_1 - 1)x_1^2\end{aligned}\quad (45)$$

Adaptive backstepping with tuning functions [11] results in

$$\begin{aligned}u_1 &= -(x_2 + x_1 - 1 + x_1^2\hat{\theta}) - (x_1 - 1) - (1 + 2x_1\hat{\theta})(x_2 + \hat{\theta}x_1^2) - x_1^2\tau - x_1\hat{\theta}_1 - x_2\hat{\theta}_2 \\ \dot{\hat{\theta}} &= \tau; \quad \tau = (x_1 - 1)x_1^2 + (x_2 + x_1 - 1 + x_1^2\hat{\theta})x_1^2(1 + 2x_1\hat{\theta}) \\ \dot{\hat{\theta}}_1 &= (x_2 + x_1 - 1 + x_1^2\hat{\theta})x_1; \quad \dot{\hat{\theta}}_2 = (x_2 + x_1 - 1 + x_1^2\hat{\theta})x_2\end{aligned}\quad (46)$$

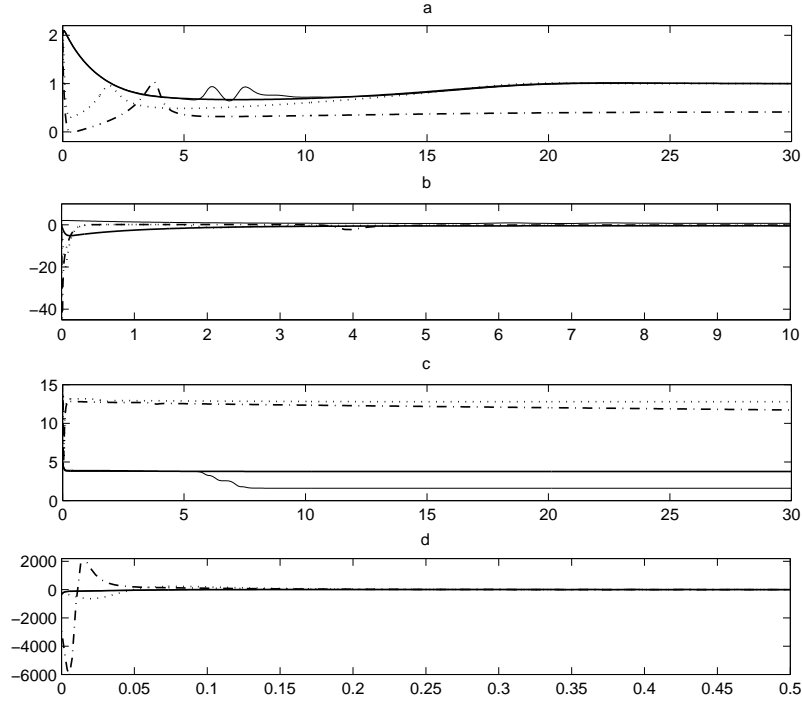


Figure 1: Plots of system (34), (47) trajectories with control functions (43),(44) (thick solid lines), (45) (dotted line), (46) (dash-dotted line), and (48) (thin solid line). Plot $a - x_1$ as a function of time, $b - x_2$ as a function of time, $c - \Delta\hat{\theta}$ as a function of time, $d - u$ as a function of time.

We simulated the adaptive system dynamics for the following set of parameters and initial conditions $x_1(0) = 2$, $x_2(0) = 0.2$, $\theta = 1$, $\theta_3(0) = \theta(0) = 3$, $\theta_1(0) = \theta_2(0) = -2$, $\xi_2(0) = 0$, $\xi_1(0) = 0$, $k = 10$. Initial conditions for $\hat{\theta}_{1,I}(0)$, $\hat{\theta}_{2,I}(0)$ and $\hat{\theta}_{3,I}(0)$ where chosen to satisfy $\hat{\theta}_1(0) = \hat{\theta}_2(0) = -2$, $\hat{\theta}_3(0) = 3$. Parameters θ_1 , θ_2 : $\theta_1 = 1$, $\theta_2 = 0.5$. As an additional measure of performance, we introduced the variable $\Delta\hat{\theta}$ which indicates the distance in the controller parameter space between the estimates and real values of the parameters. Simulation results are presented in Figure 1. In Figure 1 thick solid lines show the system dynamics with algorithm (43),(44), dotted lines show the system behavior with algorithm (45), and dash-dotted lines correspond to algorithm (46). It turns out that system (34) with algorithm (46) also reaches the goal manifold, but after around 400 seconds of modeling time. We can see again that transient performance of the adaptive algorithms proposed in the paper is better than that of conventional algorithms. In addition, we calculated the integral $I = \int_0^T u_1^2(\tau) d\tau$, $T = 500$, for every controller along the system solutions. The values of the functional I indicate how much energy is spent to achieve the control goal. For control function (43),(44) $I = 627.10$, for adaptive backstepping controller (45) $I = 13329.28$, for controller (46) $I = 263872.58$. To illustrate the ability of our algorithms to deal with nonlinear parameterization, we change (34) to

$$\dot{x}_1 = x_1^2\theta + x_2; \quad \dot{x}_2 = 5 \tanh(x_1\theta_1 + x_2\theta_2) + u_2 \quad (47)$$

Nonlinearity $\tanh(x_1\theta_1 + x_2\theta_2)$ satisfies Assumption 5 with respect to function $\alpha(\mathbf{x}) = (x_1, x_2)^T$ and, in addition,

Assumption 6 is also satisfied for any bounded x_1 and x_2 . Then according to Theorem 4, control function

$$\begin{aligned} u = & -5 \tanh(\xi \hat{\theta}_1 + x_2 \hat{\theta}_2) - \xi^2(x_1 - 1)x_1^2 - (x_2 + \xi - 1 + \frac{1}{3}\xi^5 + \hat{\theta}_{0,I}\xi^2) - \\ & (1 + \frac{5}{3}\xi^4 + 2\xi\hat{\theta}_{0,I})((x_1 - \xi)(F^2(x_1, \xi, \hat{\theta}_{0,I}) + 1) + \frac{1}{3}x_1^5 + \hat{\theta}_{\xi,I}(t)x_1^2 + x_2) \end{aligned} \quad (48)$$

along with (44) guarantees that $\psi_1, \psi_2, \dot{\psi}_1, \dot{\psi}_2 \in L_2 \cap L_\infty$. The simulation results of system (47) with control algorithm (48), (44) are given in Figure 1 (thin solid lines). The value of functional I for this case is 3186.83.

5 Conclusions

The method proposed in this paper suggests a new methodology to design adaptive control algorithms. Our method to design of the adaptation schemes is consistent with recent trends in adaptive control, for instance, [12], where nonlinear controllers are proposed to adaptively stabilize linear plants. Indeed, when derived for linear systems algorithms in finite form will also result in nonlinearities in the controller. These nonlinearities are to be introduced, in particular, to improve the performance of the adaptive system. In contrast to [12], we show not only that the L_2 and L_∞ norm bounds are computable for the state vector, but also that properties P1)–P7) are ensured. The method, however, is different from conventional approaches, as it is not restricted by realizability issues. While in conventional parametric adaptive control the realizability of adaptation schemes in differential form determines the properties of the resulting systems (including poor performance and restricted applicability), in our method we first determine the desired properties of the controller (Theorem 1, Proposition 1 and Lemmas 1 – 2) and only then deal with the realizability problem. In order to realize the adaptive algorithms in finite form explicitly, i.e. without extension of the system state space, special restrictions formulated in Assumption 7 are to be satisfied (Theorem 2 and Corollaries 1, 2).

To realize the adaptation algorithms that do not satisfy the explicit realizability conditions formulated in Assumption 7, we embed the original system into a system of higher order. This system should satisfy a-priori certain structural conditions that are formulated in Corollaries 1 and 2. These two ideas (design of an algorithm aiming for its best properties, not its realizability, and design of an embedding for realization) result in a new method, which is shown to be applicable to a sufficiently large class of systems with nonlinear parameterization, e.g., like those given by systems (32). It is very important that no damping or discontinuities are injected directly into the control function in contrast to [9, 15, 16].

In the present article we hope to have extended the scope of applicability and performance of adaptive control algorithms. Our results to date are applicable to the full-state feedback case only. Extension of the results to the output-feedback case remains a future study topic.

References

- [1] A. Astolfi and R. Ortega. Immersion and invariance: A new tool for stabilization and adaptive control of nonlinear systems. *IEEE Trans. on Automatic Control*, 48(4):590–605, 2003.

- [2] A. Datta and P. Ioannou. Directly computable l_2 and l_∞ performance bounds for Morse's dynamic certainty equivalence adaptive controller. *Int. Journal of Adaptive Control and Signal Processing*, 9:423–432, 1995.
- [3] V. Fomin, A. Fradkov, and V. Yakubovich. *Adaptive Control of Dynamical Systems*. Nauka, 1981.
- [4] A. L. Fradkov. Integro-differentiating velocity gradient algorithms. *Sov. Phys. Dokl.*, 31(2):97–98, 1986.
- [5] A. L. Fradkov. *Adaptive Control for the Complex Systems*. Nauka, 1990.
- [6] M. French. An analytical comparison between the nonsingular quadratic performance of robust and adaptive backstepping designs. *IEEE Trans. on Automatic Control*, 47(4):670–675, 2002.
- [7] M. French, Cs. Szepesvari, and E. Rogers. Uncertainty, performance, and model dependency in approximate adaptive nonlinear control. *IEEE Trans. on Automatic Control*, 45(2):353–358, 2000.
- [8] I. Kanellakopoulos, P. V. Kokotovic, and A. S. Morse. Systematic design of adaptive controllers for feedback linearizable systems. *IEEE Trans. on Automatic Control*, 36:1241–1253, 1991.
- [9] A. Kojic and A.M. Annaswamy. Adaptive Control of Nonlinearly Parameterized Systems with a Triangular Structure. *Automatica*, 38(1):115–123, 2002.
- [10] G. Kreisselmeier. Adaptive observers with exponential rate of convergence. *IEEE Trans. Automatic Control*, AC-22:2–8, 1977.
- [11] M. Krstic, I. Kanellakopoulos, and P. Kokotovic. Adaptive nonlinear control without overparametrization. *Systems and Control Letters*, 19:177–185, 1992.
- [12] M. Krstic, I. Kanellakopoulos, and P. Kokotovic. Nonlinear design of adaptive controllers for linear systems. *IEEE Trans. on Automatic Control*, 39(4):738–752, 1994.
- [13] M. Krstic, I. Kanellakopoulos, and P. Kokotovic. *Nonlinear and Adaptive Control Design*. Wiley and Sons Inc., 1995.
- [14] Tong-Heng Lee and K. S. Narendra. Robust adaptive control of discrete-time systems using persistent excitation. *Automatica*, 24(6):781–788, 1988.
- [15] W. Lin and C. Qian. Adaptive control of nonlinearly parameterized systems: A nonsmooth feedback framework. *IEEE Trans. Automatic Control*, 47(5):757–773, 2002.
- [16] W. Lin and C. Qian. Adaptive control of nonlinearly parameterized systems: The smooth feedback case. *IEEE Trans. Automatic Control*, 47(8):1249–1266, 2002.
- [17] Ai-Poh Loh, A.M. Annaswamy, and F.P. Skantze. Adaptation in the presence of general nonlinear parameterization: An error model approach. *IEEE Trans. on Automatic Control*, 44(9):1634–1652, 1999.

- [18] R. Marino and P. Tomei. Global adaptive output-feedback control of nonlinear systems, part I: Linear parameterization. *IEEE Trans. Automatic Control*, 38(1):17–32, 1993.
- [19] A.S. Morse. High-order parameter tuners for the adaptive control of linear and nonlinear systems. In *Proceedings US-Italy Workshop in honor of Professor Antonio Ruberti, Carpi, 15-17 June. 1992*.
- [20] K. S. Narendra and A. M. Annaswamy. *Stable Adaptive systems*. Prentice–Hall, 1989.
- [21] K. S. Narendra and J. Balakrishnan. Improving transient response of adaptive control systems using multiple models and switching. *IEEE Trans. on Automatic Control*, 39(9):1861–1866, 1994.
- [22] M.S. Netto, A.M. Annaswamy, R. Ortega, and P. Moya. Adaptive control of a class of nonlinearly parameterized system using convexification. *International Journal of Control*, 73(14):1312–1321, 2000.
- [23] R. Ortega. On Morse’s new adaptive controller: Parameter convergence and transient performance. *IEEE Trans. on Automatic Control*, 38(8):1191–1202, 1993.
- [24] R. Ortega, A. Astolfi, and N. E. Barabanov. Nonlinear PI Control of Uncertain Systems: an Alternative to Parameter Adaptation. *Systems and Control Letters*, 47:259–278, 2002.
- [25] D.V. Prokhorov, V.A. Terekhov, and I.Yu. Tyukin. On the applicability conditions for the algorithms of adaptive control in nonconvex problems. *Automation and Remote Control*, 63(2):262–279, 2002.
- [26] S. Sastry and M. Bodson. *Adaptive Control: Stability, Convergence, and Robustness*. Prentice Hall, 1989.
- [27] A. Stotsky. Lyapunov design for convergence rate improvement in adaptive control. *International Journal of Control*, 57(2):501–504, 1993.
- [28] I. Y. Tyukin. Algorithms in finite form for nonlinear dynamic objects. *Automation and Remote Control*, 64(6):951–974, 2003.
- [29] I. Yu. Tyukin, D. V. Prokhorov, and Cees van Leeuwen. Finite form realizations of adaptive control algorithms. In *Proceedings of IEE European Control Conference, Cambridge, UK, September 1–4. 2003*.
- [30] I.Yu. Tyukin, D.V. Prokhorov, and V.A. Terekhov. Adaptive control with nonconvex parameterization. *IEEE Trans. on Automatic Control*, 48(4):554–567, 2003.

6 Appendix 1

In this section we consider auxiliary and technical results that are used in the paper. Let the system dynamics with respect to the function $\psi(\mathbf{x}, t)$ be described as follows:

$$\dot{\psi} = -\varphi(\psi) + z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) + \varepsilon(t), \quad (49)$$

where function $\varepsilon : R_+ \rightarrow R$, $\varepsilon \in C^0$ models unknown disturbances due to unmodeled dynamics or measurement errors. In addition, we will assume that the adaptation algorithms are affected by a disturbance:

$$\dot{\boldsymbol{\theta}} = \Gamma((\dot{\psi} + \varphi(\psi(\mathbf{x}, t)))\boldsymbol{\alpha}(\mathbf{x}, t) + \delta(t)), \quad \delta : R_+ \rightarrow R^d, \quad \delta \in C^0$$

Lemma 1 *Let error model (49) be given, $\delta, \varepsilon \in L_\infty$, $|\varphi(\psi)| > K|\psi|$, $K > 0$ and Assumptions 1, 3–6 hold for $\varepsilon \equiv 0$. Then $\psi(\mathbf{x}, t)$ and $\hat{\boldsymbol{\theta}}, \mathbf{x}(t)$ are bounded for the error model (49) with algorithm*

$$\dot{\hat{\boldsymbol{\theta}}} = (\Gamma(\dot{\psi} + \varphi(\psi(\mathbf{x}, t)))\boldsymbol{\alpha}(\mathbf{x}, t) + \delta(t) - \lambda\hat{\boldsymbol{\theta}}), \quad \lambda > 0. \quad (50)$$

Lemma 1 proof. Denote $Q(\psi) = \int_0^\psi \varphi(\varsigma) d\varsigma$ and consider the following function

$$V(\psi, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}^*) = 2(D - D_1)Q(\psi) + 0.5\|\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*\|_{\Gamma^{-1}}^2. \quad (51)$$

Its derivative satisfies the following ($z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) + \varepsilon(t) = \varphi(\psi) + \dot{\psi}$ due to equation (49)):

$$\begin{aligned} \dot{V} &= -2(D - D_1)\varphi(\psi)\dot{\psi} + (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)^T((\varphi(\psi) + \dot{\psi})\boldsymbol{\alpha}(\mathbf{x}, t) + \delta(t) + \lambda\hat{\boldsymbol{\theta}}) = -2(D - D_1)\varphi\dot{\psi} - \\ &\quad (z(\mathbf{x}, \boldsymbol{\theta}, t) + z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t))(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)^T\boldsymbol{\alpha}(\mathbf{x}, t) + \varepsilon(t)(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)^T\boldsymbol{\alpha}(\mathbf{x}, t) + (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)^T(\delta(t) - \lambda\hat{\boldsymbol{\theta}}) \end{aligned}$$

From Assumptions 5, 6 it follows that

$$\begin{aligned} &-(z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) - z(\mathbf{x}, \boldsymbol{\theta}, t))(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)^T\boldsymbol{\alpha}(\mathbf{x}, t) + \varepsilon(t)(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)^T\boldsymbol{\alpha}(\mathbf{x}, t) - \frac{D_1\varepsilon^2(t)}{4} + \frac{D_1\varepsilon^2(t)}{4} \leq \\ &\leq -D(z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t))^2 + D_1|\varepsilon(t)||z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)| - \frac{D_1\varepsilon^2(t)}{4} + \frac{D_1\varepsilon^2(t)}{4} \\ &= -(D - D_1)(z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t))^2 - D_1\left(|z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)| - \frac{|\varepsilon(t)|}{2}\right)^2 + \frac{D_1\varepsilon^2(t)}{4} \end{aligned} \quad (52)$$

Then

$$\begin{aligned} \dot{V} &\leq 2(D - D_1)\varphi(\psi)\dot{\psi} - (D - D_1)(z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t))^2 - D_1\left(|z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)| - \frac{|\varepsilon(t)|}{2}\right)^2 \\ &\quad + \frac{D_1\varepsilon^2(t)}{4} - (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)^T(\delta(t) + \lambda\hat{\boldsymbol{\theta}}) \leq -(D - D_1)(\varphi^2(\psi) + \dot{\psi}^2 - 2(\varphi(\psi) + \dot{\psi})\varepsilon(t) + \varepsilon^2(t)) + \\ &\quad + \frac{D_1\varepsilon^2(t)}{4} - \lambda\|\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*\|^2 + \|\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*\|(\|\delta(t)\| + \lambda\|\hat{\boldsymbol{\theta}}^*\|) \\ &\leq -(D - D_1)\left(1 - \frac{1}{\Delta_1^2}\right)(\varphi^2(\psi) + \dot{\psi}^2) + 2(D - D_1)\Delta_1^2\varepsilon^2(t) + \frac{D_1\varepsilon^2(t)}{4} - \lambda\left(1 - \frac{1}{\Delta_2^2}\right)\|\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*\|^2 + \\ &\quad \Delta_2^2\frac{(\|\delta(t)\| + \lambda\|\hat{\boldsymbol{\theta}}^*\|)^2}{4\lambda}, \end{aligned} \quad (53)$$

where $\Delta_1, \Delta_2 > 1$. In the lemma conditions $\delta, \varepsilon \in L_\infty$. Therefore, taking into account estimate (53) and inequality $|\varphi(\psi)| > K|\psi|$, we conclude that derivative \dot{V} is negative-definite for any $\psi, \hat{\boldsymbol{\theta}}$ that belong to the following set:

$$\begin{aligned} \Omega_{t>0} &= \left\{ \psi, \hat{\boldsymbol{\theta}} \left| (D - D_1)\left(1 - \frac{1}{\Delta_1^2}\right)\varphi^2(\psi) + \lambda\left(1 - \frac{1}{\Delta_2^2}\right)\|\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*\|^2 \geq \right. \right. \\ &\quad \left. \left. \|\varepsilon(t)\|_\infty^2 \left(2(D - D_1)\Delta_1^2 + \frac{D_1}{4}\right) + \Delta_2^2\frac{(\|\delta(t)\|_\infty + \lambda\|\hat{\boldsymbol{\theta}}^*\|)^2}{4\lambda} \right\} \end{aligned}$$

Hence $\hat{\boldsymbol{\theta}}, \psi(\mathbf{x}, t)$ are bounded. *The lemma is proven.*

Let $\varepsilon \in L_2$ and $\delta(t) \equiv 0$. In this case it is possible to show that the control goal is reached in the closed loop system with slightly modified version of algorithm (12).

Lemma 2 *Let the following error model be given*

$$\dot{\psi} = -\varphi(\psi)(1 + F(t)) + z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) + \varepsilon(t), \quad (54)$$

where $F : R_+ \rightarrow R_+$, $F(t) \in C^0$, $\varepsilon(t) \in C^0$, $\varepsilon(t) \in L_2$, Assumptions 3-6 hold for $F(t) \equiv 0$, $\varepsilon(t) \equiv 0$ and adaptation algorithm satisfy equation

$$\dot{\hat{\boldsymbol{\theta}}} = \Gamma(\dot{\psi} + \varphi(\psi)(1 + F(t))\boldsymbol{\alpha}(\mathbf{x}, t). \quad (55)$$

Then

- 1) $\psi(\mathbf{x}, t) \in L_\infty$, $\varphi(\psi(\mathbf{x}, t)) \in L_2 \cap L_\infty$, $\sqrt{F(t)}\varphi(\psi(\mathbf{x}, t)) \in L_2$; $\hat{\boldsymbol{\theta}} \in L_\infty$
- 2) $z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) \in L_2$.

If $F(t) \in L_\infty$ then

- 3) $\dot{\psi} \in L_2$.

If in addition functions $\varepsilon(t) \in L_\infty$ and function $z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)$ is locally bounded with respect to \mathbf{x} , $\hat{\boldsymbol{\theta}}$, uniformly bounded with respect to t then

- 4) $\psi(\mathbf{x}, t) \rightarrow 0$ as $t \rightarrow \infty$.

Lemma 2 proof. Function $\varepsilon(t) \in L_2$, therefore integral $\int_t^\infty \varepsilon^2(\tau) d\tau < \infty$. Consider the following function

$$V_{\hat{\boldsymbol{\theta}}}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}^*, t) = \frac{D_1}{4} \int_t^\infty \varepsilon^2(\tau) d\tau + \frac{1}{2} \|\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*\|_{\Gamma^{-1}}^2.$$

Its time-derivative can be written as follows:

$$\begin{aligned} \dot{V}_{\hat{\boldsymbol{\theta}}} &= -1/4 D_1 \varepsilon^2(t) + (\varphi(\psi)(1 + F(t)) + \dot{\psi})(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)^T \boldsymbol{\alpha}(\mathbf{x}, t) \\ &= -1/4 D_1 \varepsilon^2(t) + (z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) + \varepsilon(t))(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)^T \boldsymbol{\alpha}(\mathbf{x}, t) \end{aligned}$$

Taking into account inequality (52) we can write the following estimate for $\dot{V}_{\hat{\boldsymbol{\theta}}}$:

$$\dot{V}_{\hat{\boldsymbol{\theta}}} \leq -(D - D_1)(z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t))^2 = -(D - D_1)(\varphi(\psi)(1 + F(t)) + \dot{\psi} - \varepsilon(t))^2 \quad (56)$$

It follows from (56) that $z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}^*, t) \in L_2$. Let us denote $\mu(t) = \varepsilon(t) + z(\mathbf{x}(t), \hat{\boldsymbol{\theta}}(t), t) - z(\mathbf{x}(t), \hat{\boldsymbol{\theta}}^*, t)$. Taking this equality into account, error model (54) can be written as follows $\dot{\psi} = -(1 + F(t))\varphi(\psi) + \mu(t)$, where function $\mu(t) \in L_2$ as a sum of the functions from L_2 . Consider the following nonnegative function $V_1(\psi, t)$:

$$V_1(\psi, t) = \int_0^\psi \varphi(\xi) d\xi + \frac{1}{4} \int_t^\infty \mu^2(\tau) d\tau$$

Its time-derivative is:

$$\dot{V}_1 = -\varphi(\psi)(1 + F(t))\varphi(\psi) + \varphi(\psi)\mu(t) - \frac{1}{4}\mu^2(t) - F(t)\varphi^2(\psi) - \left(\varphi(\psi) - \frac{1}{2}\mu(t)\right)^2 \quad (57)$$

It follows from inequality (57) that $\psi(\mathbf{x}, t), \varphi(\psi(\mathbf{x}(t), t)) \in L_\infty$. Furthermore, $\sqrt{F(t)}\varphi(\psi(\mathbf{x}(t), t)) \in L_2$ and $(\varphi(\psi(\mathbf{x}(t), t)) - \mu(t)/2) \in L_2$. Given that $\mu(t) \in L_2$ it is clear that $\varphi(\psi(\mathbf{x}(t), t)) \in L_2$. Hence statements 1) and 2) of the lemma are proven. Let $F(t) \in L_\infty$ then $(1 + F(t))\varphi(\psi(\mathbf{x}(t), t)) \in L_2$ and therefore according to (54) $\dot{\psi} \in L_2$ as well. Thus statement 3) is proven. To show that 4) holds it is sufficient to notice that $\hat{\theta}$ is bounded due to (56). According to Assumption 1 state \mathbf{x} is bounded as $\psi(\mathbf{x}, t)$ is bounded. Then $\dot{\psi}$ is bounded if $\varepsilon(t)$ is bounded and function $\mathbf{z}(\mathbf{x}, \theta, t)$ is locally bounded. Hence applying Barbalat's lemma we conclude that $\psi(\mathbf{x}, t) \rightarrow 0$ as $t \rightarrow \infty$. *The lemma is proven.*

Lemma 3 *Let system:*

$$\dot{x}_i = f_i(x_1, \dots, x_i, \theta_i) + \beta_i(\mathbf{x}, t), \quad i = 1, \dots, n \quad (58)$$

and smooth function $u(\mathbf{x}, \mathbf{z}, \theta_0) : R^n \times R^m \times R^d \rightarrow R$, be given. Let us assume that $\theta_0 \in \Omega_0$, Ω_0 be bounded and there exist smooth functions $\bar{F}(\mathbf{x}, \mathbf{x}', \mathbf{z}), \bar{D}_i(\mathbf{x}, \mathbf{x}'), i = 1, \dots, n$ such that the following properties hold:

$$(u(\mathbf{x}, \mathbf{z}, \theta_0) - u(\mathbf{x}', \mathbf{z}, \theta_0))^2 \leq \|\mathbf{x} - \mathbf{x}'\|^2 \bar{F}^2(\mathbf{x}, \mathbf{x}', \mathbf{z}) \quad \forall \theta_0 \in \Omega_0, \mathbf{x}, \mathbf{x}' \in R^n$$

$$(f_i(x_1, \dots, x_i, \theta_i) - f_i(x'_1, \dots, x'_i, \theta_i))^2 \leq \|\tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}'_i\|^2 \bar{D}_i^2(\mathbf{x}_i, \mathbf{x}'_i) \quad \forall \theta_i \in \Omega_\theta, \mathbf{x}_i, \mathbf{x}'_i \in R^n$$

$$\tilde{\mathbf{x}}_i = (x_1, \dots, x_i, 0, \dots, 0)^T, \quad \tilde{\mathbf{x}}'_i = (x'_1, \dots, x'_i, 0, \dots, 0)^T$$

Let us also assume that there exist $\alpha_i(\mathbf{x})$ such that Assumptions 5, 6 hold for the functions $f_i(x_1, \dots, x_i, \theta_i)$ respectively.

Then there exist $\xi(t) : R \rightarrow R^n, \nu(t) : R \rightarrow R^m$, smooth functions $\mathbf{f}_\xi(\cdot), \mathbf{f}_\nu(\cdot)$ and corresponding system:

$$\begin{aligned} \dot{\xi} &= \mathbf{f}_\xi(\mathbf{x}, \xi, \mathbf{z}, \nu), \quad \xi_0 \in R^n \\ \dot{\nu} &= \mathbf{f}_\nu(\mathbf{x}, \xi, \mathbf{z}, \nu), \quad \nu_0 \in R^m \end{aligned} \quad (59)$$

such that

$$1) u(\mathbf{x}, \mathbf{x}, \theta_0) - u(\mathbf{q}_i, \mathbf{z}, \theta_0) \in L_2, \quad i = 1, \dots, n$$

$$\mathbf{q}_i = (\xi_1, \dots, \xi_i, x_{i+1}, \dots, x_n)^T;$$

$$2) f_i(x_1, \dots, x_i, \theta_i) - f_i(\xi_1, \dots, \xi_{i-1}, x_i, \theta_i) \in L_2, \quad i = 2, \dots, n;$$

$$3) \mathbf{x} \in L_\infty \Rightarrow \xi, \nu \in L_\infty.$$

Lemma 3 proof. For the sake of notational convenience we would like to use the following notations:

$$f_i(x_1, \dots, x_i, \theta_i) = f_i(\mathbf{x}, \theta_i), \quad \psi_{\xi_i} = x_i - \xi_i$$

$$\varepsilon_i(t) = f_i(\mathbf{q}_{i-2}, \theta_i) - f_i(\mathbf{q}_{i-1}, \theta_i), \quad i = 2, \dots, n. \quad (60)$$

Consider the following system of differential equations:

$$\begin{aligned}\dot{\xi}_i &= ((\bar{F}_i^2(\mathbf{q}_{i-1}, \mathbf{q}_i, \mathbf{z}) + \sum_{j=i}^k \bar{D}_{j+1}^2(\mathbf{q}_{i-1}, \mathbf{q}_i)) + 1)(x_i - \xi_i) + f_i(\mathbf{q}_{i-1}, \hat{\boldsymbol{\theta}}_{\xi_i}) + \beta_i(\mathbf{x}, t) \\ \dot{\hat{\boldsymbol{\theta}}}_{\xi_i} &= \gamma_{\xi_i}(\psi_{\xi_i}((\bar{F}_i^2(\mathbf{q}_{i-1}, \mathbf{q}_i, \mathbf{z}) + \sum_{j=i}^k \bar{D}_{j+1}^2(\mathbf{q}_{i-1}, \mathbf{q}_i)) + 1) + \dot{\psi}_{\xi_i})\boldsymbol{\alpha}_i(\mathbf{q}_{i-1}), \gamma_{\xi_i} > 0, i = 1, \dots, k,\end{aligned}\tag{61}$$

where $\bar{F}_i(\mathbf{q}_{i-1}, \mathbf{q}_i, \mathbf{z}) = \bar{F}(\mathbf{q}_{i-1}, \mathbf{q}_i, \mathbf{z})$. Taking into account (58) and (61) let us write the following error model:

$$\begin{aligned}\dot{\psi}_{\xi_i} &= -((\bar{F}_i^2(\mathbf{q}_{i-1}, \mathbf{q}_i, \mathbf{z}) + \sum_{j=i}^k \bar{D}_{j+1}^2(\mathbf{q}_{i-1}, \mathbf{q}_i)) + 1)\psi_{\xi_i} - f_i(\mathbf{q}_{i-1}, \hat{\boldsymbol{\theta}}_{\xi_i}) + f_i(\mathbf{x}, \boldsymbol{\theta}_i) \\ \dot{\hat{\boldsymbol{\theta}}}_{\xi_i} &= \gamma_{\xi_i}(\psi_{\xi_i}((\bar{F}_i^2(\mathbf{q}_{i-1}, \mathbf{q}_i, \mathbf{z}) + \sum_{j=i}^k \bar{D}_{j+1}^2(\mathbf{q}_{i-1}, \mathbf{q}_i)) + 1) + \dot{\psi}_{\xi_i})\boldsymbol{\alpha}_i(\mathbf{q}_{i-1}), \gamma_{\xi_i} > 0\end{aligned}\tag{62}$$

It is clear that trajectories of system (61) for $k = 1$ satisfy the following condition $u(\mathbf{x}, \mathbf{z}, \boldsymbol{\theta}_0) - u(\mathbf{q}_1, \mathbf{z}, \boldsymbol{\theta}_0) \in L_2$ (this follows directly from Lemma 2). Consider the case when $k = 2$. Taking into account the equations for $\dot{\psi}_{\xi_1}$ and $\dot{\hat{\boldsymbol{\theta}}}_{\xi_1}$ we can derive from Lemma 2 that

$$\sqrt{\bar{F}_1^2(\mathbf{q}_0, \mathbf{q}_1, \mathbf{z}) + \bar{D}_2^2(\mathbf{q}_0, \mathbf{q}_1)}(x_1 - \xi_1) \in L_2, x_1 - \xi_1 \in L_\infty, \hat{\boldsymbol{\theta}}_{\xi_1} \in L_\infty$$

Hence $\bar{F}_1(\mathbf{q}_0, \mathbf{q}_1, \mathbf{z})(x_1 - \xi_1) \in L_2$, $\bar{D}_2(\mathbf{q}_0, \mathbf{q}_1)(x_1 - \xi_1) \in L_2$ as $\sqrt{\bar{F}_1^2(\mathbf{q}_0, \mathbf{q}_1, \mathbf{z}) + \bar{D}_2^2(\mathbf{q}_0, \mathbf{q}_1)} \geq |\bar{F}_1(\mathbf{q}_0, \mathbf{q}_1, \mathbf{z})|$ and $\sqrt{\bar{F}_1^2(\mathbf{q}_0, \mathbf{q}_1, \mathbf{z}) + \bar{D}_2^2(\mathbf{q}_0, \mathbf{q}_1)} \geq |\bar{D}_2(\mathbf{q}_0, \mathbf{q}_1)|$. Therefore, we can conclude that

$$u(\mathbf{q}_0, \mathbf{z}, \boldsymbol{\theta}_0) - u(\mathbf{q}_1, \mathbf{z}, \boldsymbol{\theta}_0) \in L_2, f_2(\mathbf{q}_0, \boldsymbol{\theta}_2) - f_2(\mathbf{q}_1, \boldsymbol{\theta}_2) \in L_2$$

as $|u(\mathbf{q}_0, \mathbf{z}, \boldsymbol{\theta}_0) - u(\mathbf{q}_1, \mathbf{z}, \boldsymbol{\theta}_0)| \leq |\bar{F}_1(\mathbf{q}_0, \mathbf{q}_1, \mathbf{z})(x_1 - \xi_1)| \in L_2$, $|f_2(\mathbf{q}_0, \boldsymbol{\theta}_2) - f_2(\mathbf{q}_1, \boldsymbol{\theta}_2)| \leq |\bar{D}_2(\mathbf{q}_0, \mathbf{q}_1)(x_1 - \xi_1)| \in L_2$. Notice that $\varepsilon_2(t) = f_2(\mathbf{q}_0, \boldsymbol{\theta}_2) - f_2(\mathbf{q}_1, \boldsymbol{\theta}_2) = f_2(\mathbf{x}, \boldsymbol{\theta}_2) - f_2(\mathbf{q}_1, \boldsymbol{\theta}_2)$. Therefore, the equations for $\dot{\psi}_{\xi_2}$ become as follows:

$$\begin{aligned}\dot{\psi}_{\xi_2} &= -(\bar{F}_2^2(\mathbf{q}_1, \mathbf{q}_2, \mathbf{z}) + 1)\psi_{\xi_2} - f_2(\mathbf{q}_1, \hat{\boldsymbol{\theta}}_{\xi_2}) + f_2(\mathbf{q}_1, \boldsymbol{\theta}_2) + \varepsilon_2(t) \\ \dot{\hat{\boldsymbol{\theta}}}_{\xi_2} &= \gamma_{\xi_2}(\psi_{\xi_2}(\bar{F}_2^2(\mathbf{q}_1, \mathbf{q}_2, \mathbf{z}) + 1) + \dot{\psi}_{\xi_2})\boldsymbol{\alpha}_2(\mathbf{q}_1), \gamma_{\xi_2} > 0, \varepsilon(t) \in L_2\end{aligned}\tag{63}$$

Hence, applying Lemma 2 to system (62), (63) and taking into account that $|u(\mathbf{q}_1, \mathbf{z}, \boldsymbol{\theta}_0) - u(\mathbf{q}_2, \mathbf{z}, \boldsymbol{\theta}_0)| \leq |\bar{F}_2(\mathbf{q}_1, \mathbf{q}_2, \mathbf{z})(x_2 - \xi_2)|$ we can conclude that

$$\begin{aligned}u(\mathbf{q}_0, \mathbf{z}, \boldsymbol{\theta}_0) - u(\mathbf{q}_1, \mathbf{z}, \boldsymbol{\theta}_0) &\in L_2, x_1 - \xi_1 \in L_\infty, \hat{\boldsymbol{\theta}}_{\xi_1} \in L_\infty \\ u(\mathbf{q}_1, \mathbf{z}, \boldsymbol{\theta}_0) - u(\mathbf{q}_2, \mathbf{z}, \boldsymbol{\theta}_0) &\in L_2, x_2 - \xi_2 \in L_\infty, \hat{\boldsymbol{\theta}}_{\xi_2} \in L_\infty \\ f_2(\mathbf{q}_0, \boldsymbol{\theta}_2) - f_2(\mathbf{q}_1, \boldsymbol{\theta}_2) &\in L_2,\end{aligned}$$

and, subsequently, $u(\mathbf{q}_0, \mathbf{z}, \boldsymbol{\theta}_0) - u(\mathbf{q}_2, \mathbf{z}, \boldsymbol{\theta}_0) \in L_2$ as a sum of two signals from L_2 .

Let us now consider arbitrary $2 < k \leq n$. It follows from Lemma 2 that for the error model with respect to function ψ_{ξ_1} :

$$\dot{\psi}_{\xi_1} = -((\bar{F}_1^2(\mathbf{q}_0, \mathbf{q}_1, \mathbf{z}) + \sum_{j=1}^k \bar{D}_{j+1}^2(\mathbf{q}_0, \mathbf{q}_1)) + 1)\psi_{\xi_1} - f_1(\mathbf{q}_0, \hat{\boldsymbol{\theta}}_{\xi_1}) + f_1(\mathbf{x}, \boldsymbol{\theta}_1)$$

and corresponding subsystem

$$\dot{\hat{\theta}}_{\xi_1} = \gamma_{\xi_1}(\psi_{\xi_1}((\bar{F}_1^2(\mathbf{q}_0, \mathbf{q}_1, \mathbf{z}) + \sum_{j=1}^k \bar{D}_{j+1}^2(\mathbf{q}_0, \mathbf{q}_1)) + 1) + \dot{\psi}_{\xi_1})\alpha_1(\mathbf{q}_0),$$

one can derive that $\sqrt{\bar{F}_1^2(\mathbf{q}_0, \mathbf{q}_1, \mathbf{z}) + \sum_{j=1}^k \bar{D}_{j+1}^2(\mathbf{q}_0, \mathbf{q}_1)}(x_1 - \xi_1) \in L_2$. This in consequence implies that

$$\begin{aligned} u(\mathbf{x}, \mathbf{z}, \boldsymbol{\theta}_0) - u(\mathbf{q}_1, \mathbf{z}, \boldsymbol{\theta}_0) &\in L_2 \\ f_i(\mathbf{x}, \boldsymbol{\theta}_i) - f_i(\mathbf{q}_1, \boldsymbol{\theta}_i) &\in L_2, \quad i = 2, \dots, k \end{aligned}$$

Hence we can write the error model for ψ_{ξ_2} in (62) in the following form

$$\dot{\psi}_{\xi_2} = -((\bar{F}_2^2(\mathbf{q}_1, \mathbf{q}_2, \mathbf{z}) + \sum_{j=3}^k \bar{D}_{j+1}^2(\mathbf{q}_1, \mathbf{q}_2)) + 1)\psi_{\xi_2} - f_2(\mathbf{q}_1, \hat{\boldsymbol{\theta}}_{\xi_2}) + f_2(\mathbf{q}_1, \boldsymbol{\theta}_2) + \varepsilon_2(t)$$

where $\varepsilon_2(t) \in L_2$. It follows from Lemma 2 that

$$\begin{aligned} u(\mathbf{q}_1, \mathbf{z}, \boldsymbol{\theta}_0) - u(\mathbf{q}_2, \mathbf{z}, \boldsymbol{\theta}_0) &\in L_2 \\ f_i(\mathbf{q}_1, \boldsymbol{\theta}_i) - f_i(\mathbf{q}_2, \boldsymbol{\theta}_i) &\in L_2, \quad i = 3, \dots, k \end{aligned}$$

in system (61). Notice also that $f_3(\mathbf{x}, \boldsymbol{\theta}_3) - f_3(\mathbf{q}_2, \boldsymbol{\theta}_3) \in L_2$ as a sum of two functions from L_2 . By the similar reasoning it is can be shown that for any $2 \leq i \leq n$ we can represent the error model system (62) as follows

$$\dot{\psi}_{\xi_i} = -((\bar{F}_i^2(\mathbf{q}_{i-1}, \mathbf{q}_i, \mathbf{z}) + \sum_{j=i+1}^k \bar{D}_{j+1}^2(\mathbf{q}_{i-1}, \mathbf{q}_i)) + 1)\psi_{\xi_i} - f_i(\mathbf{q}_{i-1}, \hat{\boldsymbol{\theta}}_{\xi_i}) + f_i(\mathbf{q}_{i-1}, \boldsymbol{\theta}_i) + \varepsilon_i(t),$$

where $\varepsilon_i(t) \in L_2$. Therefore, using Lemma 2 again we can conclude that

$$\begin{aligned} u(\mathbf{q}_{j-1}, \mathbf{z}, \boldsymbol{\theta}_0) - u(\mathbf{q}_j, \mathbf{z}, \boldsymbol{\theta}_0) &\in L_2 \\ f_i(\mathbf{q}_{j-1}, \boldsymbol{\theta}_i) - f_i(\mathbf{q}_j, \boldsymbol{\theta}_i) &\in L_2, \quad x_i - \xi_i \in L_\infty, \quad \hat{\boldsymbol{\theta}}_{\xi_i} \in L_\infty, \quad i = j, \dots, n, \end{aligned}$$

The last, however, implies that $u(\mathbf{x}, \mathbf{z}, \boldsymbol{\theta}_0) - u(\mathbf{q}_i, \mathbf{z}, \boldsymbol{\theta}_0) \in L_2$. In order to complete the proof we have to make sure that system (61) is physically realizable. In particular, realization of subsystems

$$\dot{\hat{\theta}}_{\xi_i} = \gamma_{\xi_i}(\psi_{\xi_i}((\bar{F}_i^2(\mathbf{q}_{i-1}, \mathbf{q}_i, \mathbf{z}) + \sum_{j=i}^k \bar{D}_{j+1}^2(\mathbf{q}_{i-1}, \mathbf{q}_i)) + 1) + \dot{\psi}_{\xi_i})\alpha_i(\mathbf{q}_{i-1}), \quad \gamma_{\xi_i} > 0, \quad i = 1, \dots, k, \quad (64)$$

shell not be dependent on any uncertainties $\boldsymbol{\theta}_i$. It follows, however, from Theorem 2 that there are realizations of algorithms (64) in finite form:

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{\xi_i}(\mathbf{q}_{i-1}, \xi_i, t) &= \gamma_{\xi_i}(\hat{\boldsymbol{\theta}}_{\xi_i, P}(\mathbf{q}_{i-1}, \xi_i) + \hat{\boldsymbol{\theta}}_{\xi_i, I}(t)), \quad \gamma_{\xi_i} > 0 \\ \hat{\boldsymbol{\theta}}_{\xi_i, P}(\mathbf{q}_{i-1}, \xi_i) &= \psi_{\xi_i}(x_i, \xi_i)\alpha_i(\mathbf{q}_{i-1}) - \Psi_{\xi_i}(\mathbf{q}_{i-1}, \xi_i) \\ \dot{\hat{\boldsymbol{\theta}}}_{\xi_i, I} &= ((\bar{F}_i^2(\mathbf{q}_{i-1}, \mathbf{q}_i, \mathbf{z}) + \sum_{j=i}^k \bar{D}_{j+1}^2(\mathbf{q}_{i-1}, \mathbf{q}_i)) + 1)\psi_{\xi_i}(x_i, \xi_i)\alpha_i(\mathbf{q}_{i-1}) + \\ &\quad \sum_{j=1}^i \frac{\partial \Psi_{\xi_i}(\mathbf{q}_{i-1}, \xi_i)}{\partial \xi_j} \dot{\xi}_j - \sum_{j=1}^{i-1} \psi_{\xi_i}(x_i, \xi_i) \frac{\partial \alpha_i(\mathbf{q}_{i-1})}{\partial \xi_j} \dot{\xi}_j \\ \Psi_{\xi_i}(\mathbf{q}_{i-1}, \xi_i) &= \int_{x_i(0)}^{x_i(t)} \psi_{\xi_i}(x_i, \xi_i) \frac{\partial \alpha_i(\mathbf{q}_{i-1})}{\partial x_i} dx_i \end{aligned} \quad (65)$$

Notice also that if $\mathbf{x} \in L_\infty$ then $\boldsymbol{\xi} \in L_\infty$ and hence $\hat{\boldsymbol{\theta}}_{\xi_i, P}(\mathbf{q}_{i-1}, \xi_i) \in L_\infty$ as $\hat{\boldsymbol{\theta}}_{\xi_i, P}(\mathbf{q}_{i-1}, \xi_i)$ is smooth. Given that $\hat{\boldsymbol{\theta}}_{\xi_i} = \gamma_{\xi_i}(\hat{\boldsymbol{\theta}}_{\xi_i, P}(\mathbf{q}_{i-1}, \xi_i) + \hat{\boldsymbol{\theta}}_{\xi_i, I})$ and both $\hat{\boldsymbol{\theta}}_{\xi_i}, \hat{\boldsymbol{\theta}}_{\xi_i, P}(\mathbf{q}_{i-1}, \xi_i) \in L_\infty$ then we can conclude that $\hat{\boldsymbol{\theta}}_{\xi_i, I} \in L_\infty$ for $\mathbf{x} \in L_\infty$.

It is easy to see that denoting $\boldsymbol{\nu} = \hat{\boldsymbol{\theta}}_{\xi_i, I}$ we can transform system (61), (65) into (59) which satisfies statements 1)–3) of the lemma. *The lemma is proven.*

7 Appendix 2

Theorem 1 proof. Let us consider the following positive-definite function: $V_{\hat{\boldsymbol{\theta}}}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}^*) = \frac{1}{2}\|\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*\|_{\Gamma^{-1}}^2$. Its time-derivative according to equations (13) can be derived as follows: $\dot{V}_{\hat{\boldsymbol{\theta}}}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}^*) = (\varphi(\psi) + \dot{\psi})(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)^T \boldsymbol{\alpha}(\mathbf{x}, t)$. According to Assumption 5 and equality (9) it is easy to see that

$$\dot{V}_{\hat{\boldsymbol{\theta}}}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}^*) = -(z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) - z(\mathbf{x}, \boldsymbol{\theta}, t))(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)^T \boldsymbol{\alpha}(\mathbf{x}, t) \leq -D(z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) - z(\mathbf{x}, \boldsymbol{\theta}, t))^2 = -D(\varphi(\psi) + \dot{\psi})^2 \leq 0 \quad (66)$$

Therefore $V_{\hat{\boldsymbol{\theta}}}$ is non-increasing (property P2) is proven). Furthermore, integration of $\dot{V}_{\hat{\boldsymbol{\theta}}}$ with respect to time results in

$$V_{\hat{\boldsymbol{\theta}}}(\hat{\boldsymbol{\theta}}(0), \hat{\boldsymbol{\theta}}^*) - V_{\hat{\boldsymbol{\theta}}}(\hat{\boldsymbol{\theta}}(t), \hat{\boldsymbol{\theta}}^*) \geq D \int_0^t (\dot{\psi}(\tau) + \varphi(\psi(\tau)))^2 d\tau \geq 0.$$

Function $V_{\hat{\boldsymbol{\theta}}}$ is non-increasing and bounded from below as $V_{\hat{\boldsymbol{\theta}}} \geq 0$, therefore

$$D \int_0^t (\dot{\psi}(\tau) + \varphi(\psi(\tau)))^2 d\tau \leq V_{\hat{\boldsymbol{\theta}}}(\hat{\boldsymbol{\theta}}(0), \hat{\boldsymbol{\theta}}^*) < \infty.$$

Hence $(\varphi(\psi) + \dot{\psi}) = (z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)) = (z(\mathbf{x}, \hat{\boldsymbol{\theta}}^*, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)) \in L_2$ (property P3)).

To prove property P1) let us consider the following function: $V(\psi, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}^*) = 2DQ(\psi) + V_{\hat{\boldsymbol{\theta}}}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}^*)$, where $Q(\psi) = \int_0^\psi \varphi(\varsigma) d\varsigma$. Function $V(\psi, \hat{\boldsymbol{\theta}})$ is positive-definite with respect to $\psi(\mathbf{x}, t)$ and $\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*$ because of Assumption 4. Its time-derivative obeys inequality: $\dot{V}(\psi, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}^*) \leq 2D\varphi(\psi)\dot{\psi} - D(\dot{\psi} + \varphi(\psi))^2 = -D\varphi^2(\psi) - D\dot{\psi}^2 \leq 0$.

Therefore, function $V(\psi, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}^*)$ is bounded and non-increasing. Furthermore

$$\begin{aligned} \infty &> V(\psi(\mathbf{x}(0), 0), \hat{\boldsymbol{\theta}}(0), \hat{\boldsymbol{\theta}}^*) \geq V(\psi(\mathbf{x}(0), 0), \hat{\boldsymbol{\theta}}(0), \hat{\boldsymbol{\theta}}^*) - V(\psi(\mathbf{x}(t), t), \hat{\boldsymbol{\theta}}(t), \hat{\boldsymbol{\theta}}^*) \geq D \int_0^t \varphi^2(\psi(\mathbf{x}(\tau), \tau)) d\tau \geq 0 \\ \infty &> V(\psi(\mathbf{x}(0), 0), \hat{\boldsymbol{\theta}}(0), \hat{\boldsymbol{\theta}}^*) \geq V(\psi(\mathbf{x}(0), 0), \hat{\boldsymbol{\theta}}(0), \hat{\boldsymbol{\theta}}^*) - V(\psi(\mathbf{x}(t), t), \hat{\boldsymbol{\theta}}(t), \hat{\boldsymbol{\theta}}^*) \geq D \int_0^t \dot{\psi}^2(\tau) d\tau \geq 0. \end{aligned} \quad (67)$$

or, equivalently, $\dot{\psi}(t) \in L_2$, $\varphi(\psi(t)) \in L_2$. Hence, property P1) is proven as well. The L_2 norm bounds (14) for $\varphi(\psi)$ and $\dot{\psi}$ follow immediately from inequality (67):

$$\|\varphi(\psi)\|_2^2 \leq D^{-1}V(\psi(\mathbf{x}(0), 0), \hat{\boldsymbol{\theta}}(0), \hat{\boldsymbol{\theta}}^*), \quad \|\dot{\psi}\|_2^2 \leq D^{-1}V(\psi(\mathbf{x}(0), 0), \hat{\boldsymbol{\theta}}(0), \hat{\boldsymbol{\theta}}^*)$$

The L_∞ norm bound for $\psi(\mathbf{x}(t), t)$ results from the inequality: $V(\psi(\mathbf{x}(0), 0), \hat{\boldsymbol{\theta}}(0), \hat{\boldsymbol{\theta}}^*) - V(\psi(\mathbf{x}(t), t), \hat{\boldsymbol{\theta}}(t), \hat{\boldsymbol{\theta}}^*) \geq 0$. Consider function Λ defined as $\Lambda(d) = \max_{|\psi|} \{|\psi| \mid \int_0^{|\psi|} \varphi(\varsigma) d\varsigma = d\}$ and notice that it is monotonic and nondecreasing. Therefore, given that $\int_0^{\psi(\mathbf{x}(t), t)} \varphi(\varsigma) d\varsigma \leq \frac{1}{2D}V(\psi(\mathbf{x}(0), 0), \hat{\boldsymbol{\theta}}(0), \hat{\boldsymbol{\theta}}^*)$ we can conclude that $|\psi| \leq \Lambda\left(\frac{1}{2D}V(\psi(\mathbf{x}(0), 0), \hat{\boldsymbol{\theta}}(0), \hat{\boldsymbol{\theta}}^*)\right)$. To prove property P4) notice that function $V(\psi(\mathbf{x}(t), t), \hat{\boldsymbol{\theta}}(t), \hat{\boldsymbol{\theta}}^*)$ is

bounded. Hence by Assumption 4 function $\psi(\mathbf{x}(t), t)$ is bounded as well. According to Assumption 1 boundedness of $\psi(\mathbf{x}(t), t)$ implies boundedness of the state \mathbf{x} . In addition it is assumed that $z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)$ is locally bounded with respect to $\mathbf{x}, \hat{\boldsymbol{\theta}}$ and uniformly bounded in t . Therefore the difference $z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)$ is bounded. Furthermore, by Assumption 4 function $\varphi(\psi) \in C^0$ and therefore it is bounded as well given that ψ is bounded. Hence $\dot{\psi}$ is bounded and by applying Barbalat's lemma one can show that $\psi(\mathbf{x}(t), t) \rightarrow 0$ at $t \rightarrow \infty$.

To complete the proof of the theorem consider the difference $z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)$. Let function $\varphi \in C^1$, function $z(\mathbf{x}, \boldsymbol{\theta}, t)$ is differentiable in $\mathbf{x}, \boldsymbol{\theta}$; derivative $\partial z(\mathbf{x}, \boldsymbol{\theta}, t)/\partial t$ is bounded uniformly in t ; function $\boldsymbol{\alpha}(\mathbf{x}, t)$ is locally bounded with respect to \mathbf{x} and uniformly bounded with respect to t , then $d/dt(z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t))$ is bounded. On the other hand there exists the following limit

$$\lim_{t \rightarrow \infty} \int_0^t (z(\mathbf{x}, \boldsymbol{\theta}, \tau) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, \tau))^2 d\tau = \int_0^\infty (z(\mathbf{x}, \boldsymbol{\theta}, \tau) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, \tau))^2 d\tau \leq \frac{1}{D} V_{\hat{\boldsymbol{\theta}}}(\hat{\boldsymbol{\theta}}(0), \hat{\boldsymbol{\theta}}^*)$$

as $\int_0^t (z(\mathbf{x}, \boldsymbol{\theta}, \tau) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, \tau))^2 d\tau$ is non-decreasing and bounded from above. Hence by Barbalat's lemma it follows that $z(\mathbf{x}, \boldsymbol{\theta}, \tau) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, \tau) \rightarrow 0$ as $t \rightarrow \infty$. Notice also that $\psi(\mathbf{x}(t), t) \rightarrow 0$ as $t \rightarrow \infty$. Then $\dot{\psi} \rightarrow 0$ as $t \rightarrow \infty$. *The theorem is proven.*

Proof of Proposition 1. Consider the following integral⁷ $\int_0^t (\dot{\psi}(\tau) + \varphi(\psi(\tau)))^2 d\tau$. It was shown in Theorem 1 proof that $\int_0^t (\dot{\psi}(\tau) + \varphi(\psi(\tau)))^2 d\tau \leq \frac{1}{2D} \|\hat{\boldsymbol{\theta}}(0) - \hat{\boldsymbol{\theta}}^*\|_{\Gamma^{-1}}^2$ along system (13) solutions. Let us define $\mu(t) = \dot{\psi}(t) + \varphi(\psi(t))$. In the other words

$$\dot{\psi} = -\varphi(\psi) + \mu(t), \quad (68)$$

where $\int_0^\infty \mu^2(\tau) d\tau \leq \frac{1}{2D} \|\hat{\boldsymbol{\theta}}(0) - \hat{\boldsymbol{\theta}}^*\|_{\Gamma^{-1}}^2$. According to the proposition conditions, $\varphi(\psi) = K\psi$, it is possible to derive the solution of equation (68) as follows $\psi(t) = \psi(0)e^{-Kt} + \int_0^t e^{-K(t-\tau)} \mu(\tau) d\tau$. Hence

$$\begin{aligned} |\psi(t)| &\leq |\psi(0)|e^{-Kt} + \sqrt{\left(\int_0^t e^{-K(t-\tau)} \mu(\tau) d\tau\right)^2} \leq |\psi(0)|e^{-Kt} + \sqrt{\int_0^t e^{-2K(t-\tau)} d\tau \int_0^t \mu^2(\tau) d\tau} \\ &\leq |\psi(0)|e^{-Kt} + \frac{1}{2} \sqrt{\frac{1}{KD} \|\hat{\boldsymbol{\theta}}(0) - \hat{\boldsymbol{\theta}}^*\|_{\Gamma^{-1}}^2}. \end{aligned} \quad (69)$$

Property P6) is thus proven. In order to prove property P7) consider

$$\dot{\hat{\boldsymbol{\theta}}} = \Gamma(\dot{\psi} + \varphi(\psi))\boldsymbol{\alpha}(\mathbf{x}, t) = \Gamma(z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t))\boldsymbol{\alpha}(\mathbf{x}, t).$$

Function

$$\begin{aligned} D_1 |\boldsymbol{\alpha}(\mathbf{x}, t)^T (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)| &\leq |z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)| \leq D |\boldsymbol{\alpha}(\mathbf{x}, t)^T (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*)| \\ \boldsymbol{\alpha}(\mathbf{x}, t)^T (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*) (z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}^*, t)) &> 0 \quad \forall \quad z(\mathbf{x}, \hat{\boldsymbol{\theta}}^*, t) \neq z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t). \end{aligned}$$

Therefore, there exists $D_1 \leq \kappa(t) \leq D$ such that

$$\dot{\hat{\boldsymbol{\theta}}} = -\kappa(t) \Gamma \boldsymbol{\alpha}(\mathbf{x}, t)^T (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*) \boldsymbol{\alpha}(\mathbf{x}, t) = -\kappa(t) \Gamma \boldsymbol{\alpha}(\mathbf{x}, t) \boldsymbol{\alpha}(\mathbf{x}, t)^T (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*).$$

⁷That we substitute the arguments of the functions $\dot{\psi}(\cdot)$ and $\psi(\cdot)$ with t means that we consider them as functions of time.

Hence

$$\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}^* = e^{-\Gamma \int_0^t \kappa(\tau) \boldsymbol{\alpha}(\mathbf{x}(\tau), \tau) \boldsymbol{\alpha}(\mathbf{x}(\tau), \tau)^T d\tau} (\hat{\boldsymbol{\theta}}(0) - \hat{\boldsymbol{\theta}}^*) \quad (70)$$

Consider the integral $\Gamma \int_0^t \kappa(\tau) \boldsymbol{\alpha}(\mathbf{x}(\tau), \tau) \boldsymbol{\alpha}(\mathbf{x}(\tau), \tau)^T d\tau$ for $t > L$

$$\Gamma \int_0^t \kappa(\tau) \boldsymbol{\alpha}(\mathbf{x}(\tau), \tau) \boldsymbol{\alpha}(\mathbf{x}(\tau), \tau)^T d\tau \geq \Gamma D_1 \int_0^t \boldsymbol{\alpha}(\mathbf{x}(\tau), \tau) \boldsymbol{\alpha}(\mathbf{x}(\tau), \tau)^T d\tau,$$

where $\boldsymbol{\alpha}(\mathbf{x}(t), t)$ is persistently exciting. For any $t > L$ there exists integer $n \geq 0$ such that $t = nL + r$, $r \in R, 0 \leq r < L$. Therefore

$$\Gamma D_1 \int_0^t \boldsymbol{\alpha}(\mathbf{x}(\tau), \tau) \boldsymbol{\alpha}(\mathbf{x}(\tau), \tau)^T d\tau \geq \Gamma D_1 n \delta I \geq \left(\frac{\Gamma D_1 \delta}{L} t - I \right).$$

Then taking into account (70) one can write

$$\|\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}^*\| \leq \|e^{(-\frac{\Gamma D_1 \delta}{L} t + I)}\| \|\hat{\boldsymbol{\theta}}(0) - \hat{\boldsymbol{\theta}}^*\|, \quad (71)$$

i. e. $\hat{\boldsymbol{\theta}}(t)$ converges to $\hat{\boldsymbol{\theta}}^*$ exponentially fast. It means that there exist positive constants $\lambda > 0$, $\lambda \neq K$ and $D_{\hat{\boldsymbol{\theta}}} > 0$ such that $\|\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}^*\| \leq e^{-\lambda t} \|\hat{\boldsymbol{\theta}}(0) - \hat{\boldsymbol{\theta}}^*\| D_{\hat{\boldsymbol{\theta}}}$. It follows from Theorem 1 that $\psi(\mathbf{x}(t), t)$ is bounded. In addition due to Assumption 1 we can conclude that \mathbf{x} is bounded as well. By the proposition assumptions function $\boldsymbol{\alpha}(\mathbf{x}, t)$ is locally bounded with respect to \mathbf{x} and uniformly bounded in t . Therefore, there exists $D_{\boldsymbol{\alpha}} > 0$ such that $|\boldsymbol{\alpha}(\mathbf{x}, t)^T (\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}^*)| \leq D_{\boldsymbol{\alpha}} \|\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}^*\|$. Taking into account that $z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t) = \mu(t)$ and $|z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)| \leq D |\boldsymbol{\alpha}(\mathbf{x}, t)^T (\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}^*)|$ one can derive from (68) the following estimate

$$|\psi(t)| \leq |\psi(0)| e^{-Kt} + \|\hat{\boldsymbol{\theta}}(0) - \hat{\boldsymbol{\theta}}^*\| D_{\hat{\boldsymbol{\theta}}} D_{\boldsymbol{\alpha}} D \int_0^t e^{-K(t-\tau)} e^{-\lambda \tau} d\tau \leq |\psi(0)| e^{-Kt} + \frac{D_{\hat{\boldsymbol{\theta}}} D_{\boldsymbol{\alpha}} D}{K - \lambda} \|\hat{\boldsymbol{\theta}}(0) - \hat{\boldsymbol{\theta}}^*\| e^{-\lambda t} \quad (72)$$

The proposition is proven.

Proof of Theorem 2. The theorem proof is quite straightforward and follows from explicit differentiation of function $\hat{\boldsymbol{\theta}}(\mathbf{x}, t)$ with respect to time: $\dot{\hat{\boldsymbol{\theta}}}(\mathbf{x}, t) = \Gamma(\dot{\boldsymbol{\theta}}_P + \dot{\boldsymbol{\theta}}_I) = \Gamma(\dot{\psi} \boldsymbol{\alpha}(\mathbf{x}, t) + \psi \dot{\boldsymbol{\alpha}}(\mathbf{x}, t) - \dot{\Psi}(\mathbf{x}, t) + \dot{\boldsymbol{\theta}}_I)$. Notice that

$$\begin{aligned} \psi \dot{\boldsymbol{\alpha}}(\mathbf{x}, t) - \dot{\Psi}(\mathbf{x}, t) + \dot{\boldsymbol{\theta}}_I &= \psi(\mathbf{x}, t) \frac{\partial \boldsymbol{\alpha}(\mathbf{x}, t)}{\partial \mathbf{x}_1} \dot{\mathbf{x}}_1 + \psi(\mathbf{x}, t) \frac{\partial \boldsymbol{\alpha}(\mathbf{x}, t)}{\partial \mathbf{x}_2} \dot{\mathbf{x}}_2 + \psi(\mathbf{x}, t) \frac{\partial \boldsymbol{\alpha}(\mathbf{x}, t)}{\partial t} - \\ &\frac{\partial \Psi(\mathbf{x}, t)}{\partial \mathbf{x}_1} \dot{\mathbf{x}}_1 - \frac{\partial \Psi(\mathbf{x}, t)}{\partial \mathbf{x}_2} \dot{\mathbf{x}}_2 - \frac{\partial \Psi(\mathbf{x}, t)}{\partial t} + \dot{\boldsymbol{\theta}}_I \end{aligned} \quad (73)$$

According to Assumption 7, $\frac{\partial \Psi(\mathbf{x}, t)}{\partial \mathbf{x}_2} = \psi(\mathbf{x}, t) \frac{\partial \boldsymbol{\alpha}(\mathbf{x}, t)}{\partial \mathbf{x}_2}$. Then taking into account (73), we can obtain

$$\psi \dot{\boldsymbol{\alpha}}(\mathbf{x}, t) - \dot{\Psi}(\mathbf{x}, t) + \dot{\boldsymbol{\theta}}_I = \left(\psi(\mathbf{x}, t) \frac{\partial \boldsymbol{\alpha}(\mathbf{x}, t)}{\partial \mathbf{x}_1} - \frac{\partial \Psi}{\partial \mathbf{x}_1} \right) \dot{\mathbf{x}}_1 + \psi(\mathbf{x}, t) \frac{\partial \boldsymbol{\alpha}(\mathbf{x}, t)}{\partial t} - \frac{\Psi(\mathbf{x}, t)}{\partial t} + \dot{\boldsymbol{\theta}}_I \quad (74)$$

Notice that according to the proposed notations we can rewrite the term $\left(\psi(\mathbf{x}, t) \frac{\partial \boldsymbol{\alpha}(\mathbf{x}, t)}{\partial \mathbf{x}_1} - \frac{\partial \Psi}{\partial \mathbf{x}_1} \right) \dot{\mathbf{x}}_1$ in the following form: $(\psi(\mathbf{x}, t) L_{\mathbf{f}_1} \boldsymbol{\alpha}(\mathbf{x}, t) - L_{\mathbf{f}_1} \Psi(\mathbf{x}, t)) + (\psi(\mathbf{x}, t) L_{\mathbf{g}_1} \boldsymbol{\alpha}(\mathbf{x}, t) - L_{\mathbf{g}_1} \Psi(\mathbf{x}, t)) u(\mathbf{x}, \hat{\boldsymbol{\theta}}, t)$. Hence it follows from (17) and (74) that $\psi \dot{\boldsymbol{\alpha}}(\mathbf{x}, t) - \dot{\Psi}(\mathbf{x}, t) + \dot{\boldsymbol{\theta}}_I = \varphi(\psi) \boldsymbol{\alpha}(\mathbf{x}, t)$. Therefore $\dot{\hat{\boldsymbol{\theta}}}(\mathbf{x}, t) = \Gamma(\dot{\psi} + \varphi(\psi)) \boldsymbol{\alpha}(\mathbf{x}, t)$. The theorem is proven.

Proof of Theorem 3. To prove the theorem, first notice that control function (29) provides the following error model dynamics

$$\dot{\psi} = -\varphi(\psi) + z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \hat{\boldsymbol{\theta}}, t), \quad z(\tilde{\mathbf{x}}, \hat{\boldsymbol{\theta}}, t) = L_{\nu(\tilde{\mathbf{x}}, \hat{\boldsymbol{\theta}})} \psi(\tilde{\mathbf{x}}, t). \quad (75)$$

By adding and subtracting the function $z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t)$ from the right-hand side of (75) we get the following:

$$\dot{\psi} = -\varphi(\psi) + z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t) + z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \hat{\boldsymbol{\theta}}, t),$$

where the difference $z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t)$ is bounded due to Assumption 9. Denote $\varepsilon(t) = z(\mathbf{x}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t)$, then

$$\dot{\psi} = -\varphi(\psi) + \varepsilon(t) + z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t) - z(\tilde{\mathbf{x}}, \hat{\boldsymbol{\theta}}, t), \quad (76)$$

where $\varepsilon \in L_\infty$. Denote

$$\begin{aligned} \tilde{\mathbf{f}} &= \mathbf{f}_1(\mathbf{x}) \oplus \mathbf{f}_2'(\mathbf{x}) \oplus \frac{\partial \mathbf{h}_\xi}{\partial \boldsymbol{\xi}} \mathbf{f}_\xi(\mathbf{x}, \boldsymbol{\xi}, t) \\ \tilde{\mathbf{g}} &= \mathbf{g}_1(\mathbf{x}) \oplus \mathbf{g}_2'(\mathbf{x}) \oplus \mathbf{0}_{\mathbf{x}_2''}, \quad \mathbf{0}_{\mathbf{x}_2''} = \underbrace{(0, \dots, 0)^T}_{\dim \mathbf{x}_2''}. \end{aligned}$$

Let us consider the following adaptation algorithm:

$$\begin{aligned} \hat{\boldsymbol{\theta}}(\mathbf{x}, \tilde{\mathbf{x}}, t) &= \Gamma(\hat{\boldsymbol{\theta}}_P(\mathbf{x}, \tilde{\mathbf{x}}, t) + \hat{\boldsymbol{\theta}}_I(t)), \quad \Gamma > 0; \quad \hat{\boldsymbol{\theta}}_P(\mathbf{x}, \tilde{\mathbf{x}}, t) = \psi(\mathbf{x}, t)\boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) - \Psi(\tilde{\mathbf{x}}, t); \\ \dot{\hat{\boldsymbol{\theta}}}_I &= \varphi(\psi(\mathbf{x}, t))\boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) + \frac{\partial \Psi(\tilde{\mathbf{x}}, t)}{\partial t} - \psi(\mathbf{x}, t)\frac{\partial \boldsymbol{\alpha}(\tilde{\mathbf{x}}, t)}{\partial t} - (\psi(\mathbf{x}, t)L_{\tilde{\mathbf{f}}}\boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) - L_{\tilde{\mathbf{f}}}\Psi(\tilde{\mathbf{x}}, t)) - \\ &\quad (\psi(\mathbf{x}, t)L_{\tilde{\mathbf{g}}}\boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) - L_{\tilde{\mathbf{g}}}\Psi(\tilde{\mathbf{x}}, t))u(\mathbf{x}, \mathbf{h}_\xi, \hat{\boldsymbol{\theta}}, t) - \lambda\hat{\boldsymbol{\theta}}(\tilde{\mathbf{x}}, t), \end{aligned} \quad (77)$$

where $\lambda > 0$. Differentiation of function $\hat{\boldsymbol{\theta}}_P$ with respect to time leads to:

$$\begin{aligned} \dot{\hat{\boldsymbol{\theta}}}_P(\mathbf{x}, \tilde{\mathbf{x}}, t) &= \dot{\psi}(\mathbf{x}, t)\boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) + \psi(\mathbf{x}, t)\dot{\boldsymbol{\alpha}}(\tilde{\mathbf{x}}, t) - \dot{\Psi}(\tilde{\mathbf{x}}, t) = \dot{\psi}(\mathbf{x}, t)\boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) + \psi(\mathbf{x}, t)\frac{\partial \boldsymbol{\alpha}(\tilde{\mathbf{x}}, t)}{\partial t} - \frac{\partial \Psi(\tilde{\mathbf{x}}, t)}{\partial t} \\ &\quad + \left(\psi(\mathbf{x}, t)\frac{\partial \boldsymbol{\alpha}(\tilde{\mathbf{x}}, t)}{\partial \mathbf{x}_1} - \frac{\partial \Psi(\tilde{\mathbf{x}}, t)}{\partial \mathbf{x}_1} \right) \dot{\mathbf{x}}_1 + \left(\psi(\mathbf{x}, t)\frac{\partial \boldsymbol{\alpha}(\tilde{\mathbf{x}}, t)}{\partial \mathbf{x}_2'} - \frac{\partial \Psi(\tilde{\mathbf{x}}, t)}{\partial \mathbf{x}_2'} \right) \dot{\mathbf{x}}_2' + \\ &\quad \left(\psi(\mathbf{x}, t)\frac{\partial \boldsymbol{\alpha}(\tilde{\mathbf{x}}, t)}{\partial \mathbf{h}_\xi} - \frac{\partial \Psi(\tilde{\mathbf{x}}, t)}{\partial \mathbf{h}_\xi} \right) \dot{\mathbf{h}}_\xi \end{aligned} \quad (78)$$

Taking into account (77) and (23) we can rewrite (78) as follows:

$$\begin{aligned} \dot{\hat{\boldsymbol{\theta}}}_P(\mathbf{x}, \tilde{\mathbf{x}}, t) &= \dot{\psi}(\mathbf{x}, t)\boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) - \frac{\partial \Psi(\tilde{\mathbf{x}}, t)}{\partial t} + \psi(\mathbf{x}, t)\frac{\partial \boldsymbol{\alpha}(\tilde{\mathbf{x}}, t)}{\partial t} + (\psi(\mathbf{x}, t)L_{\tilde{\mathbf{f}}}\boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) - L_{\tilde{\mathbf{f}}}\Psi(\tilde{\mathbf{x}}, t)) + \\ &\quad (\psi(\mathbf{x}, t)L_{\tilde{\mathbf{g}}}\boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) - L_{\tilde{\mathbf{g}}}\Psi(\tilde{\mathbf{x}}, t))u(\mathbf{x}, \mathbf{h}_\xi, \hat{\boldsymbol{\theta}}, t) + \left(\psi(\mathbf{x}, t)\frac{\partial \boldsymbol{\alpha}(\tilde{\mathbf{x}}, t)}{\partial \mathbf{x}_2'} - \frac{\partial \Psi(\tilde{\mathbf{x}}, t)}{\partial \mathbf{x}_2'} \right) \boldsymbol{\nu}'(\mathbf{x}, \boldsymbol{\theta}) \end{aligned} \quad (79)$$

Notice also that according to Assumption 8:

$$\frac{\partial \Psi(\tilde{\mathbf{x}}, t)}{\partial \mathbf{x}_2'} = \psi(\tilde{\mathbf{x}}, t)\frac{\partial \boldsymbol{\alpha}(\tilde{\mathbf{x}}, t)}{\partial \mathbf{x}_2'}$$

and $(\psi(\mathbf{x}, t) - \psi(\tilde{\mathbf{x}}, t))L_{\boldsymbol{\nu}'(\mathbf{x}, \boldsymbol{\theta})}\boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) \in L_\infty$ due to Assumption 9. Denoting $(\psi(\mathbf{x}, t) - \psi(\tilde{\mathbf{x}}, t))L_{\boldsymbol{\nu}'(\mathbf{x}, \boldsymbol{\theta})}\boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) = \delta(t)$ and using equalities (79) and (77) we can derive that

$$\dot{\hat{\boldsymbol{\theta}}} = \Gamma(\dot{\hat{\boldsymbol{\theta}}}_P + \dot{\hat{\boldsymbol{\theta}}}_I) = \Gamma((\dot{\psi} + \varphi(\psi(\mathbf{x}, t)))\boldsymbol{\alpha}(\tilde{\mathbf{x}}, t) + \delta(t) - \lambda\hat{\boldsymbol{\theta}}), \quad (80)$$

where function $\delta(t)$ is bounded. Let us define the extended state space vector $\mathbf{q} = \mathbf{x} \oplus \boldsymbol{\xi}$. Furthermore, we define $z_{\mathbf{q}}(\mathbf{q}, \boldsymbol{\theta}, t) = z(\tilde{\mathbf{x}}, \boldsymbol{\theta}, t)$, $\boldsymbol{\alpha}_{\mathbf{q}}(\mathbf{q}, t) = \boldsymbol{\alpha}(\tilde{\mathbf{x}}, t)$, $\psi_{\mathbf{q}}(\mathbf{q}, t) = \psi(\mathbf{x}, t)$. Given the chosen notations, Algorithm (80)

can be written as follows: $\dot{\hat{\theta}} = \Gamma((\dot{\psi}_{\mathbf{q}} + \varphi(\psi_{\mathbf{q}}))\alpha_{\mathbf{q}}(\mathbf{q}, t) + \delta(t) - \lambda\hat{\theta})$. Moreover instead of equation (76) we can write $\dot{\psi} = -\varphi(\psi) + z_{\mathbf{q}}(\mathbf{q}, \theta, t) - z_{\mathbf{q}}(\mathbf{q}, \hat{\theta}, t) + \varepsilon(t)$.

It is easy to see that Assumptions 5 and 6 hold for the extended system. Assumption 1 is also satisfied with respect to the goal function $\psi_{\mathbf{q}}(\mathbf{q}, t)$ due to hypothesis (26) in Assumption 8. Indeed, $\psi_{\mathbf{q}}(\mathbf{q}, t) = \psi(\mathbf{x}, t) \in L_{\infty} \Rightarrow \mathbf{x} \in L_{\infty} \Rightarrow \xi \in L_{\infty} \Rightarrow \mathbf{q} \in L_{\infty}$. Therefore, according to Assumption 9 and Lemma 1, we can conclude that $\psi(\mathbf{x}, t)$ is bounded and furthermore trajectories \mathbf{x}, ξ are bounded as well. Thus property P8) is proven. To prove property P9) it is sufficient to notice that $\delta(t) = 0$ either due to the equality $\partial\alpha(\tilde{\mathbf{x}}, t)/\partial\mathbf{x}'_2 \equiv 0$ or $\psi(\mathbf{x}, t) = \psi(\tilde{\mathbf{x}}, t)$. Let $\partial\alpha(\tilde{\mathbf{x}}, t)/\partial\mathbf{x}'_2 \equiv 0$, then P9) follows explicitly from Assumption 10 and Lemma 2 applied to (76) with algorithm

$$\begin{aligned}\hat{\theta}(\mathbf{x}, \tilde{\mathbf{x}}, t) &= \Gamma(\hat{\theta}_P(\mathbf{x}, \tilde{\mathbf{x}}, t) + \hat{\theta}_I(t)), \quad \Gamma > 0; \quad \hat{\theta}_P(\mathbf{x}, \tilde{\mathbf{x}}, t) = \psi(\mathbf{x}, t)\alpha(\tilde{\mathbf{x}}, t); \\ \dot{\hat{\theta}}_I &= \varphi(\psi(\mathbf{x}, t))\alpha(\tilde{\mathbf{x}}, t) - \psi(\mathbf{x}, t)\partial\alpha(\tilde{\mathbf{x}}, t)/\partial t - \psi(\mathbf{x}, t)L_{\tilde{\mathbf{f}}}\alpha(\tilde{\mathbf{x}}, t) - (\psi(\mathbf{x}, t)L_{\tilde{\mathbf{g}}}\alpha(\tilde{\mathbf{x}}, t))u(\mathbf{x}, \mathbf{h}_{\xi}, \hat{\theta}, t).\end{aligned}\quad (81)$$

which is in fact algorithm (82) for $\lambda = 0$ and $\Psi(\tilde{\mathbf{x}}, t) \equiv 0$. If $\psi(\mathbf{x}, t) = \psi(\tilde{\mathbf{x}}, t)$, then according to Lemma 2, algorithm (77) with $\lambda = 0$:

$$\begin{aligned}\hat{\theta}(\mathbf{x}, \tilde{\mathbf{x}}, t) &= \Gamma(\hat{\theta}_P(\mathbf{x}, \tilde{\mathbf{x}}, t) + \hat{\theta}_I(t)), \quad \Gamma > 0, \quad \hat{\theta}_P(\mathbf{x}, \tilde{\mathbf{x}}, t) = \psi(\mathbf{x}, t)\alpha(\tilde{\mathbf{x}}, t) - \Psi(\tilde{\mathbf{x}}, t) \\ \dot{\hat{\theta}}_I &= \varphi(\psi(\mathbf{x}, t))\alpha(\tilde{\mathbf{x}}, t) + \partial\Psi(\tilde{\mathbf{x}}, t)/\partial t - \psi(\mathbf{x}, t)\partial\alpha(\tilde{\mathbf{x}}, t)/\partial t - (\psi(\mathbf{x}, t)L_{\tilde{\mathbf{f}}}\alpha(\tilde{\mathbf{x}}, t) - L_{\tilde{\mathbf{f}}}\Psi(\tilde{\mathbf{x}}, t)) - \\ &\quad (\psi(\mathbf{x}, t)L_{\tilde{\mathbf{g}}}\alpha(\tilde{\mathbf{x}}, t) - L_{\tilde{\mathbf{g}}}\Psi(\tilde{\mathbf{x}}, t))u(\mathbf{x}, \mathbf{h}_{\xi}, \hat{\theta}, t).\end{aligned}\quad (82)$$

ensures P9) as well. *The theorem is proven.*

Proof of Theorem 4. We will prove the theorem by induction from order 1 to n for system (32). According to the theorem conditions functions $f_1(x_1, \theta_1)$, $\alpha_1(x_1)$ are smooth and therefore according to Lemma 2 and Theorem 2, there exists smooth function $u(x_1, \hat{\theta}_1)$:

$$\begin{aligned}u(x_1, \hat{\theta}_1) &= -f_1(x_1, \hat{\theta}_1) - \varphi_1(\psi(x_1)), \quad \hat{\theta}_1 = \gamma_1(\hat{\theta}_{1,P}(x_1) + \hat{\theta}_{1,I}(t)), \quad \gamma_1 > 0 \\ \hat{\theta}_{1,P}(x_1) &= \psi(x_1)\alpha_1(x_1) - \Psi(x_1), \quad \Psi(x_1) = \int_{x_1(0)}^{x_1(t)} \psi(x_1)\frac{\partial\alpha_1(x_1)}{\partial x_1}dx_1 \\ \dot{\hat{\theta}}_{1,I} &= \varphi_1(\psi(x_1))\alpha_1(x_1)\end{aligned}$$

such that $\psi_1(x_1, t) \in L_2 \cap L_{\infty}$, $\dot{\psi}_1 \in L_2$ for the system of the following type: $\dot{x}_1 = f_1(x_1, \theta_1) + u(x_1, \hat{\theta}_1) + \varepsilon_1(t)$, $\varepsilon_1(t) \in L_2$. Hence, the basis of induction is proven.

Let us assume that the theorem statements hold true for the systems of order i , i.e. there exists such smooth function $u_i(\mathbf{x}_i, \hat{\theta}_i, \xi_i, \nu_i)$, $\mathbf{x}_i, \xi_i \in R^i$, $\mathbf{x}_i = (x_1, \dots, x_i)^T$, $\xi_i = (\xi_1, \dots, \xi_i)^T$ and the corresponding goal functions $\psi_j(x_j, t)$, $j = 1, \dots, i$ such that $\psi_j(x_j, t) \in L_2 \cap L_{\infty}$, $\dot{\psi}_j \in L_2$ for system (32) of order i :

$$\begin{aligned}\dot{x}_j &= f_j(x_1, \dots, x_j, \theta_j) + x_{j+1}, \quad j \in \{1, \dots, i-1\}, \\ \dot{x}_i &= f_i(x_1, \dots, x_i, \theta_i) + u_i + \varepsilon_i(t), \quad \varepsilon_i(t) \in L_2.\end{aligned}\quad (83)$$

Therefore, in order to prove the theorem it is enough to show that its statements hold for system (32) of order $i+1$ given that it holds for the systems like (83).

According to the inductive assumption function $u_i(\mathbf{x}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\xi}_i, \boldsymbol{\nu}_i)$ is smooth. Then by Hadamar's lemma there exists such $F(\mathbf{x}_i, \mathbf{x}'_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\xi}_i, \boldsymbol{\nu}_i)$ that $u_i(\mathbf{x}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) - u_i(\mathbf{x}'_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) = F(\mathbf{x}_i, \mathbf{x}'_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\xi}_i, \boldsymbol{\nu}_i)(\mathbf{x}_i - \mathbf{x}'_i)$. Let us denote $\bar{F}_{i+1}^2(\mathbf{x}_i, \mathbf{x}'_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) = \|F(\mathbf{x}_i, \mathbf{x}'_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\xi}_i, \boldsymbol{\nu}_i)\|^2$. Furthermore, due to the theorem conditions functions $f_j(\mathbf{x}_j, \boldsymbol{\theta}_j)$, $j = 1, \dots, i+1$ satisfy the following additional assumptions:

$$(f_j(\mathbf{x}_j, \boldsymbol{\theta}_j) - f_j(\mathbf{x}'_j, \boldsymbol{\theta}_j))^2 \leq \|\mathbf{x}_j - \mathbf{x}'_j\|^2 \bar{D}_j^2(\mathbf{x}_j, \mathbf{x}'_j) \quad \forall \boldsymbol{\theta}_j \in \Omega_{\boldsymbol{\theta}}$$

Therefore, it follows from Lemma 3 that there exists a system of differential equations

$$\begin{aligned} \dot{\boldsymbol{\xi}}_{i+1} &= \mathbf{f}_{\boldsymbol{\xi}_{i+1}}(\boldsymbol{\xi}_{i+1}, \mathbf{x}, \mathbf{z}, \boldsymbol{\nu}_{i+1}), \boldsymbol{\xi} \in R^i \\ \dot{\boldsymbol{\nu}}_{i+1} &= \mathbf{f}_{\boldsymbol{\nu}_{i+1}}(\boldsymbol{\xi}_{i+1}, \mathbf{x}, \mathbf{z}), \quad \mathbf{z} = \hat{\boldsymbol{\theta}}_i \oplus \boldsymbol{\xi}_i \oplus \boldsymbol{\nu}_i \end{aligned}$$

such that

$$\begin{aligned} u_i(\mathbf{x}_i, \hat{\boldsymbol{\theta}}_i(\mathbf{x}_i, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) - u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) &\in L_2 \\ f_{i+1}(\mathbf{x}_{i+1}, \boldsymbol{\theta}_{i+1}) - f_{i+1}(\boldsymbol{\xi}_{i+1} \oplus x_{i+1}, \boldsymbol{\theta}_{i+1}) &\in L_2, \end{aligned}$$

Let us introduce new goal function $\psi_{i+1}(x_{i+1}, \boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \boldsymbol{\nu}_i, \hat{\boldsymbol{\theta}}_{I,i}) = x_{i+1} - u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i)$ and consider its time-derivative $\dot{\psi}_{i+1}$:

$$\begin{aligned} \dot{\psi}_{i+1} &= f_{i+1}(\mathbf{x}_{i+1}, \boldsymbol{\theta}_{i+1}) + u_{i+1} - L_{\mathbf{f}_{\boldsymbol{\xi}_i}} u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) - \\ &\quad L_{\mathbf{f}_{\boldsymbol{\nu}_i}} u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) - L_{\mathbf{f}_{\boldsymbol{\xi}_{i+1}}} u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) - \\ &\quad L_{\mathbf{f}_{\hat{\boldsymbol{\theta}}_i}} u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i). \end{aligned} \quad (84)$$

Denote $\varepsilon_{i+1}(t) = f_{i+1}(\mathbf{x}_{i+1}, \boldsymbol{\theta}_{i+1}) - f_{i+1}(\boldsymbol{\xi}_{i+1} \oplus x_{i+1}, \boldsymbol{\theta}_{i+1})$ and rewrite (84) in the following way:

$$\begin{aligned} \dot{\psi}_{i+1} &= f_{i+1}(\boldsymbol{\xi}_{i+1} \oplus x_{i+1}, \boldsymbol{\theta}_{i+1}) + u_{i+1} - L_{\mathbf{f}_{\boldsymbol{\xi}_i}} u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) - \\ &\quad L_{\mathbf{f}_{\boldsymbol{\nu}_i}} u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) - L_{\mathbf{f}_{\boldsymbol{\xi}_{i+1}}} u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) - \\ &\quad L_{\mathbf{f}_{\hat{\boldsymbol{\theta}}_i}} u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) + \varepsilon_{i+1}(t). \end{aligned} \quad (85)$$

Let us select input u_{i+1} as follows

$$\begin{aligned} u_{i+1} &= -\varphi_{i+1}(\psi_{i+1}(x_{i+1}, \boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \boldsymbol{\nu}_i, \hat{\boldsymbol{\theta}}_{I,i})) + L_{\mathbf{f}_{\boldsymbol{\xi}_i}} u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) + \\ &\quad L_{\mathbf{f}_{\boldsymbol{\nu}_i}} u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) + L_{\mathbf{f}_{\boldsymbol{\xi}_{i+1}}} u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) + \\ &\quad L_{\mathbf{f}_{\hat{\boldsymbol{\theta}}_i}} u_i(\boldsymbol{\xi}_{i+1}, \hat{\boldsymbol{\theta}}_i(\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \hat{\boldsymbol{\theta}}_{I,i}), \boldsymbol{\xi}_i, \boldsymbol{\nu}_i) - f_{i+1}(\boldsymbol{\xi}_{i+1} \oplus x_{i+1}, \hat{\boldsymbol{\theta}}_{i+1}) \end{aligned} \quad (86)$$

Denoting $\psi_{i+1}(x_{i+1}, \boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \boldsymbol{\nu}_i, \hat{\boldsymbol{\theta}}_{I,i}) = \psi_{i+1}(x_{i+1}, t)$ (as $\boldsymbol{\xi}_{i+1}, \boldsymbol{\xi}_i, \boldsymbol{\nu}_i, \hat{\boldsymbol{\theta}}_{I,i}$ are functions of time t) and substituting (86) and (85) into (84) we can write the following expression for $\dot{\psi}_{i+1}$

$$\dot{\psi}_{i+1} = -\varphi_{i+1}(\psi_{i+1}(x_{i+1}, t)) + f_{i+1}(\boldsymbol{\xi}_{i+1} \oplus x_{i+1}, \boldsymbol{\theta}_{i+1}) - f_{i+1}(\boldsymbol{\xi}_{i+1} \oplus x_{i+1}, \hat{\boldsymbol{\theta}}_{i+1}) + \varepsilon_{i+1}(t) \quad (87)$$

It follows from the theorem conditions that there exists such function $\boldsymbol{\alpha}_{i+1}(\boldsymbol{\xi}_{i+1} \oplus x_{i+1})$ that Assumptions 5 and 6 are satisfied for the function $f_{i+1}(\boldsymbol{\xi}_{i+1} \oplus x_{i+1}, \boldsymbol{\theta}_{i+1})$. Consider the following adaptation algorithm:

$$\dot{\hat{\boldsymbol{\theta}}}_{i+1} = \gamma_{i+1}(\dot{\psi}_{i+1} + \varphi_{i+1}(\psi_{i+1}(x_{i+1}, t)))\boldsymbol{\alpha}_{i+1}(\boldsymbol{\xi}_{i+1} \oplus x_{i+1}), \quad \gamma_{i+1} > 0 \quad (88)$$

Realization of algorithms (88) is guaranteed by Theorem 2 and can be given as follows:

$$\begin{aligned}
\hat{\theta}_{i+1}(\xi_{i+1} \oplus x_{i+1}, t) &= \gamma_{i+1}(\hat{\theta}_{i+1,P}(\xi_{i+1} \oplus x_{i+1}, t) + \hat{\theta}_{i+1,I}(t)), \quad \gamma_{i+1} > 0 \\
\hat{\theta}_{i+1,P}(\xi_{i+1} \oplus x_{i+1}, t) &= \psi_{i+1}(x_{i+1}, t)\alpha_{i+1}(\xi_{i+1} \oplus x_{i+1}) - \Psi_{i+1}(\xi_{i+1} \oplus x_{i+1}, t) \\
\dot{\hat{\theta}}_{i+1,I} &= \varphi_{i+1}(\psi(x_{i+1}, t))\alpha_{i+1}(\xi_{i+1} \oplus x_{i+1}) - L_{f_\xi}\alpha_{i+1}(\xi_{i+1} \oplus x_{i+1}) + \\
&\quad L_{f_\xi}\Psi_{i+1}(\xi_{i+1} \oplus x_{i+1}, t) + \frac{\partial \Psi_{i+1}(\xi_{i+1} \oplus x_{i+1}, t)}{\partial t} \\
\Psi_{i+1}(\xi_{i+1} \oplus x_{i+1}, t) &= \int_{x_{i+1}(0)}^{x_{i+1}(t)} \psi_{i+1}(x_{i+1}, t) \frac{\partial \alpha_{i+1}(\xi_{i+1} \oplus x_{i+1})}{\partial x_{i+1}} dx_{i+1}.
\end{aligned} \tag{89}$$

It follows from Lemma 2 that for the error model (87) with adaptation algorithm (88) and its realization (89) the following statements hold true: $\hat{\theta}_{i+1}(\xi_{i+1} \oplus x_{i+1}, t) \in L_\infty$, $\dot{\psi}_{i+1} \in L_2$, $\varphi_{i+1}(\psi_{i+1}(x_{i+1}, t)) \in L_2 \cap L_\infty$. Given that $\varepsilon_{i+1}(t) \in L_2$ we can conclude that

$$u_{i+1}(\mathbf{x}_i, \hat{\theta}_{i+1}, \xi_{i+1}, \xi_i, \nu_{i+1}, \nu_i, \hat{\theta}_{I,i}) - u_{i+1}(\mathbf{x}_i, \theta_{i+1}, \xi_{i+1}, \xi_i, \nu_{i+1}, \nu_i, \hat{\theta}_{I,i}) \in L_2$$

Let us denote $u_{i+1}(\mathbf{x}_i, \hat{\theta}_{i+1}, \xi_{i+1}, \tilde{\nu}_{i+1}) = u_{i+1}(\mathbf{x}_i, \hat{\theta}_{i+1}, \xi_{i+1}, \xi_i, \nu_{i+1}, \nu_i, \hat{\theta}_{I,i})$, $\tilde{\nu}_{i+1} = \xi_i \oplus \nu_{i+1} \oplus \nu_i \oplus \hat{\theta}_{I,i}$. According to the introduced notations it is easy to see that statement 2) of the theorem holds. In addition to this notice that the choice of appropriate function $\varphi_{i+1}(\cdot)$ in (86) is up to the designer. Therefore choosing $\varphi_{i+1}(\cdot) : |\varphi_{i+1}(\cdot)| \geq k_{i+1}|\psi_{i+1}|$, $k_{i+1} > 0$ we can guarantee that $\psi_{i+1}(x_{i+1}, t) \in L_2 \cap L_\infty$.

The last, however, according to the inductive hypothesis implies that $\psi_k(x_k, t) \in L_2 \cap L_\infty$ and $\dot{\psi}_k \in L_2$ for any $k = 1, \dots, i$, $\psi \in L_2 \cap L_\infty$ and $\dot{\psi} \in L_2$. Hence statement 1) of the theorem is proven as well.

Let us prove statement 3). According to the inductive hypothesis $\mathbf{x}_i, \xi_i, \nu_i, \hat{\theta}_i$ are bounded. Furthermore, $\hat{\theta}_{I,i}(t)$ is bounded as $\hat{\theta}_{P,i}(\mathbf{x}_i, \xi_i)$ is smooth function and $\hat{\theta}_i = \gamma_i(\hat{\theta}_{P,i}(\mathbf{x}_i, \xi_i) + \hat{\theta}_{I,i})$. Then taking into account Lemma 3 we can conclude that ξ_{i+1}, ν_{i+1} are bounded. Hence $\tilde{\nu}_{i+1}$ is bounded. Let us show that x_{i+1} is bounded as well. First notice that the difference $\varepsilon_i(t) = u_i(\mathbf{x}_i, \hat{\theta}_i(\mathbf{x}_i, \xi_i, \hat{\theta}_{I,i}), \xi_i, \nu_i) - u_i(\xi_{i+1}, \hat{\theta}_i(\xi_{i+1}, \xi_i, \hat{\theta}_{I,i}), \xi_i, \nu_i)$ is bounded as u_i is smooth and its arguments are bounded. On the other hand we have just shown that $\psi_{i+1} = x_{i+1} - u_i(\xi_{i+1}, \hat{\theta}_i(\xi_{i+1}, \xi_i, \hat{\theta}_{I,i}), \xi_i, \nu_i)$ is bounded. Therefore, x_{i+1} is bounded. Hence statement 3) is proven.

Derivatives $\dot{\psi}_j$, $j = 1, \dots, i$ are bounded as $\varepsilon_i(t)$ is bounded (according to the inductive hypothesis the theorem holds for any $j = 1 \dots, i$). If, however, $\varepsilon(t)$ is bounded then $\dot{\psi}_{i+1}$ is bounded as well as $u_{i+1}(\cdot)$, $f_{i+1}(\cdot)$ are smooth and $\mathbf{x}_{i+1}, \xi_{i+1}, \nu_{i+1}, \hat{\theta}_{i+1}$ are bounded. Therefore, it follows from Lemma 2 that $\psi_{i+1} \rightarrow 0$ as $t \rightarrow \infty$. Thus statement 4) is proven. *The theorem is proven.*